

ON VISCOSITY SOLUTION OF HJB EQUATIONS WITH STATE CONSTRAINTS AND REFLECTION CONTROL

ANUP BISWAS, HITOSHI ISHII, SUBHAMAY SAHA, AND LIN WANG

ABSTRACT. Motivated by a control problem of a certain queueing network we consider a control problem where the dynamics is constrained in the nonnegative orthant \mathbb{R}_+^d of the d -dimensional Euclidean space and controlled by the reflections at the faces/boundaries. We define a discounted value function associated to this problem and show that the value function is a viscosity solution to a certain HJB equation in \mathbb{R}_+^d with nonlinear Neumann type boundary condition. Under certain conditions, we also characterize this value function as the unique solution to this HJB equation.

CONTENTS

1. Introduction	1
2. Setting and main result	3
2.1. Motivation from queueing network	5
3. Proof of Theorem 2.1(a)	7
4. Proof of Theorem 2.1(b)	15
5. Construction of a test function	18
5.1. A convex body	18
5.2. Sign of \mathcal{H}_i	22
5.3. Construction of $\varphi(x)$	24
Appendix A.	28
Acknowledgement	31
References	32

1. INTRODUCTION

The chief goal of this article is to characterize the value function for discounted control problem associated to certain constrained dynamics (see (2.1)) as a viscosity solution of certain Hamilton-Jacobi-Bellman (HJB) equation with nonlinear Neumann type boundary condition. The controlled diffusion takes values in \mathbb{R}_+^d , non-negative orthant of the Euclidean space \mathbb{R}^d . The control process is matrix valued and represents the reflection vectors at certain time. This problem is motivated by queueing network problem where there are d

Date: November 29, 2016.

2000 Mathematics Subject Classification. 93E20, 60H30, 35J25, 49L25.

Key words and phrases. Skorokhod map with reflection control, queues with help, viscosity solutions, nonlinear Neumann boundary, non-smooth domain, heavy-traffic, stochastic network.

classes of customers in the queueing system with d number of servers, for every server there is a priority class of customers, and customers from a non-priority class may access the server provided its priority class is empty at that instant (see Section 2.1 for more details). One of the interesting features of this model is its control process which acts through reflection. To the authors knowledge such model has not been studied before. For existence and uniqueness results for such dynamics we follow [14]. The constrained dynamics in [14] does not have any control component and in fact, the author assumes certain regularity properties on the reflection matrix in the time variable. We show in Theorem A.1 that such regularity is not needed to establish existence of a unique adapted solution.

The value function is defined to be the infimum of a certain discounted cost function (see (2.4)) where the infimum is taken over all *admissible* controls. Admissible controls are processes that do not predict the future. Discounted cost functions are quite common in queueing theory (see [1] and references therein). The discounted cost considered here has two components (see (2.3)), the running cost $\ell \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ and the cost $h_i \in \mathcal{C}(\mathbb{R}^d \otimes \mathbb{R}^d)$, with $i \in \mathcal{I} := \{1, \dots, d\}$, associated to the reflections. We show that the value function is in $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ and is a viscosity solution to the following HJB equation (see (2.5))

$$\begin{cases} \mathcal{L}V - \beta V + \ell = 0 & \text{in } (\mathbb{R}_+^d)^\circ, \\ \mathcal{H}_i(DV(x)) = 0 & \text{for } x \in \partial\mathbb{R}_+^d, i \in \mathbf{I}(x), \end{cases} \quad (1.1)$$

where

$$\mathcal{H}_i(p) := \min_{M \in \mathbf{M}} \{p \cdot M e_i + h_i(M)\}, \quad i \in \mathcal{I},$$

\mathbf{M} is a certain subset of $\mathbb{R}^d \otimes \mathbb{R}^d$, and $\mathbf{I}(x)$ denotes the collection of indices $i \in \mathcal{I}$ for which $x_i = 0$. We refer [6] for more details on viscosity solutions. We note that the boundary condition here is of nonlinear Neumann type. The major hurdles in obtaining the uniqueness of viscosity solution are the non-smooth property of the boundary and the nonlinear nature of the boundary condition. The authors of [2, 7, 8] study viscosity solution in a non-smooth domain but with linear Neumann type boundary condition. In [5] viscosity solution framework is used to characterize the value function of an ergodic control problem with dynamics given by controlled constrained diffusions in a polyhedral domain. Nonlinear Neumann boundary condition is studied in [3, 9] (see also the references therein) for domains with regular boundary (at least \mathcal{C}^1). The key idea in proving uniqueness of viscosity solution is to construct a suitable test function. We could establish the uniqueness under some assumptions on h_i (see Theorem 2.1(b)). The construction of test function also uses the structure of nonlinear boundary condition.

The organization of this article is as follows. Section 2 introduces the control problem settings and the main results. Then in Section 3 we establish regularity properties of the value function and prove that the value function is a viscosity solution to (1.1), while the uniqueness of a viscosity solution to (1.1) is established in Section 4 under the assumption of the existence of a right test function. Section 5 is devoted to establishing the existence of a test function. Finally, some auxiliary results are proved in the Appendix.

Notations: By \mathbb{R}^d we denote the d -dimensional Euclidean space equipped with the Euclidean norm $|\cdot|$. By $\{e_1, \dots, e_d\}$ we denote the standard orthonormal basis in \mathbb{R}^d . We write \mathcal{I} for the set $\{1, \dots, d\}$ and set $\mathbb{R}_+^d = \{x = (x_1, \dots, x_n) \in \mathbb{R}^d : x_i \geq 0 \text{ for all } i \in \mathcal{I}\}$.

We denote $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}_+^d$. We denote by $\mathbb{R}^d \otimes \mathbb{R}^d$ the set of all $d \times d$ real matrices and we endow this space with the usual metric. For $a, b \in \mathbb{R}$ we denote the maximum (minimum) of a and b as $a \vee b$ ($a \wedge b$, respectively). We define $a^+ = a \vee 0$ and $a^- = -a \wedge 0$. Trace of a matrix $A \in \mathbb{R}^d \otimes \mathbb{R}^d$ is denoted by $\text{Tr } A$. For $x \in \mathbb{R}_+^d$ we define $\mathbb{B}_r(x) = \{y \in \mathbb{R}_+^d : |y - x| < r\}$ for $r > 0$. Given a topological space \mathcal{X} and $B \subset \mathcal{X}$, the interior, complement and boundary of B in \mathcal{X} is denoted by B°, B^c and ∂B , respectively. By $\mathcal{B}(\mathcal{X})$ we denote the Borel σ -field of \mathcal{X} . Let $\mathcal{C}([0, \infty) : \mathcal{X})$ be the set of all continuous functions from $[0, \infty)$ to \mathcal{X} . By $\mathcal{C}_0^\uparrow([0, \infty) : \mathbb{R}_+^d)$ we denote the set of all continuous paths that are non-decreasing (component-wise) with initial value 0. We define $\mathcal{C}^2(\mathbb{R}^d)$ as the set of all real valued twice continuously differentiable functions on \mathbb{R}^d . $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ denotes the set of all real valued continuous functions f with at most polynomial growth i.e.,

$$\limsup_{|x| \rightarrow 0} \frac{|f(x)|}{|x|^k} = 0 \quad \text{for some } k \geq 0.$$

Infimum over empty set is regarded as $+\infty$. $C, \kappa_1, \kappa_2, \dots$, are deterministic positive constants whose values might change from line to line.

2. SETTING AND MAIN RESULT

We consider the problem in the domain \mathbb{R}_+^d . Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which a filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfying the usual hypothesis is given. Let (W_t, \mathcal{F}_t) be a d -dimensional standard Wiener process (Brownian motion) on the above probability space, i.e., the Wiener process W_t is adapted to \mathcal{F}_t , $t \geq 0$, and for every $t, s \geq 0$, $W_{t+s} - W_t$ is independent of \mathcal{F}_t . Let $b : \mathbb{R}_+^d \rightarrow \mathbb{R}^d$ be a mapping and Σ be a $d \times d$ matrix satisfying the following condition.

Condition 2.1. There exists a $K \in (0, \infty)$ such that

$$|b(x) - b(y)| \leq K|x - y| \quad \text{and} \quad |b(x)| \leq K \quad \text{for all } x, y \in \mathbb{R}_+^d.$$

Also the matrix $A = \Sigma \Sigma^T$ is non-singular.

We note that Condition 2.1 in particular, implies that Σ is non-singular. Let $\alpha \in [0, 1]^d$. We define

$$\mathbf{M}(\alpha) := \{M = [\mathbb{I} - P] \in \mathbb{R}^d \otimes \mathbb{R}^d : P_{ii} = 0, P_{ij} \geq 0 \text{ and } \sum_{j \neq i} P_{ji} \leq \alpha_i\},$$

where \mathbb{I} denotes the $d \times d$ identity matrix. We endow that set $\mathbf{M}(\alpha)$ with the metric of $\mathbb{R}^d \otimes \mathbb{R}^d$ and therefore $\mathbf{M}(\alpha)$ forms a compact metric space. We assume that

Condition 2.2.

$$\max_i \alpha_i < 1.$$

In what follows, α is assumed to be fixed and now onwards we write $\mathbf{M}(\alpha)$ as \mathbf{M} . Given $x \in \mathbb{R}_+^d$ we consider the following reflected stochastic differential equation

$$\begin{aligned} X(t) &= x + \int_0^t b(X(s)) ds + \Sigma W(t) + \int_0^t v(s) dY(s), \quad t \geq 0, \\ \int_0^t X_i(s) dY_i(s) &= 0 \quad \text{for } i \in \mathcal{I}, \quad X \geq 0, Y \geq 0, t \geq 0, \end{aligned} \tag{2.1}$$

where the process v takes values in the compact metric space \mathbf{M} , and $v(s, \omega)$ is jointly measurable in (s, ω) . We also assume that the control process v is $\{\mathcal{F}_t\}$ adapted. It is easy to see that v is *non-anticipative* : For $s < t$, $W(t) - W(s)$ is independent of $\sigma\{W(r), v(r) : r \leq s\}$ as $W(t) - W(s)$ is independent of \mathcal{F}_s . We call the control v an *admissible* control and the set of all admissible controls is denoted by \mathfrak{M} . We show in Theorem A.1 in Appendix that if Conditions 2.1 and 2.2 hold then for every $v \in \mathfrak{M}$, (2.1) has a unique adapted strong solution (X, Y) taking values in $\mathcal{C}([0, \infty) : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, \infty) : \mathbb{R}_+^d)$.

It is well known that for every non-decreasing continuous function $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, we can define a σ -finite measure $d\phi$ on $([0, \infty), \mathcal{B}([0, \infty)))$ where $d\phi((a, b]) = \phi(b) - \phi(a)$ for $0 \leq a \leq b < \infty$. For any $f \in L^1([0, \infty), \mathcal{B}([0, \infty)), d\phi)$ we denote $\int_{[0, t]} f(s) d\phi(s) = \int_0^t f(s) d\phi(s)$. Let ℓ be a non-negative continuous function with polynomial growth, i.e., there exists $m, c_\ell \in [0, \infty)$ such that

$$0 \leq \ell(x) \leq c_\ell (1 + |x|^m), \quad x \in \mathbb{R}_+^d. \quad (2.2)$$

Let $h_i : \mathbf{M} \rightarrow \mathbb{R}, i \in \mathcal{I}$, be continuous. Since \mathbf{M} is compact, $h_i, i \in \mathcal{I}$, are bounded. Note that h_i can as well be negative. Then for any initial data $x \in \mathbb{R}_+^d$ and control $v \in \mathfrak{M}$, our cost function is given by

$$J(x, v) := \mathbb{E}_x \left[\int_0^\infty e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v(s)) dY_i(s)) \right], \quad (2.3)$$

where $\beta > 0$, and (X, Y) is the unique adapted solution to (2.1). We define the value function as

$$V(x) := \inf_{v \in \mathfrak{M}} J(x, v). \quad (2.4)$$

It is shown in Lemma 3.1 below that $V(x)$ is finite for all $x \in \mathbb{R}_+^d$. The goal of this article is to characterize the value function V as a viscosity solution of a suitable Hamilton-Jacobi-Bellman (HJB) equation. To define the HJB equation let us first introduce the following notations,

$$\mathcal{L} := \frac{1}{2} \text{Tr}(AD^2) + b \cdot D,$$

where $A = \Sigma \Sigma^T$ and D and D^2 are the gradient and Hessian operators, respectively. Define for $p \in \mathbb{R}^d$,

$$\mathcal{H}_i(p) := \min_{M \in \mathbf{M}} \{p \cdot M e_i + h_i(M)\}, \quad i \in \mathcal{I}.$$

For $x \in \mathbb{R}_+^d$ we define $\mathbf{I}(x) := \{i \in \mathcal{I} : x_i = 0\}$. Therefore $\mathbf{I}(x)$ denotes the indices of faces $F_i := \{x \in \mathbb{R}_+^d : x_i = 0\}$ where x lies. Note that $\mathbf{I}(x)$ can also be empty. Indeed, for $x \in \mathbb{R}_+^d$,

$$\mathbf{I}(x) \begin{cases} = \emptyset & \text{if } x \in (\mathbb{R}_+^d)^\circ, \\ \neq \emptyset & \text{if } x \in \partial \mathbb{R}_+^d. \end{cases}$$

The HJB equation we are interested in is given as follows

$$\begin{cases} \mathcal{L}V - \beta V + \ell = 0 & \text{in } (\mathbb{R}_+^d)^\circ, \\ \mathcal{H}_i(DV(x)) = 0 & \text{for } x \in \partial \mathbb{R}_+^d, i \in \mathbf{I}(x), \end{cases} \quad (2.5)$$

The solution of (2.5) is defined in the viscosity sense ([6–8]) as follows:

Definition 2.1 (Viscosity solution). Let $V : \mathbb{R}_+^d \rightarrow \mathbb{R}$ be a continuous function. Then V is said to be a sub (super) solution of (2.5), whenever $\varphi \in \mathcal{C}^2(\mathbb{R}^d)$ and $V - \varphi$ has a local maximum (minimum) at $x \in \mathbb{R}_+^d$, the following hold:

$$\begin{aligned} & (\mathcal{L}\varphi(x) - \beta V(x) + \ell(x)) \vee \max_{i \in \mathbf{I}(x)} \mathcal{H}_i(D\varphi(x)) \geq 0, \\ & \left((\mathcal{L}\varphi(x) - \beta V(x) + \ell(x)) \wedge \min_{i \in \mathbf{I}(x)} \mathcal{H}_i(D\varphi(x)) \leq 0, \text{ respectively} \right). \end{aligned}$$

We will impose the following condition on h_i .

Condition 2.3. For any $p \in \mathbb{R}^d$ and $x \in \partial\mathbb{R}_+^d$ if we have $\max_{i \in \mathbf{I}(x)} \mathcal{H}_i(p) < \eta$ for some $\eta \in (-\infty, 0)$ then

$$\cap_{i \in \mathbf{I}(x)} \{M \in \mathbf{M} : p \cdot Me_i + h_i(M) < \eta/2\} \neq \emptyset.$$

Remark 2.1. Condition 2.3 is assumed purely for technical reasons. This condition will be used to show that V in (2.4) is a subsolution to (2.5). It is easy to see that Condition 2.3 is satisfied if $h_i(M) \equiv h_i(M_{1i}, \dots, M_{di})$.

Our main result characterizes the value function V given by (2.4) as the viscosity solution of (2.5). Recall that $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ denotes the set of all continuous functions, defined on \mathbb{R}_+^d , with polynomial growth. Our main result is the following.

Theorem 2.1. *Let Conditions 2.1, 2.2 and 2.3 hold. Then the following holds.*

- (a) *The value function V given by (2.4) is a viscosity solution to the HJB given by (2.5);*
- (b) *Assume that $h_i(M) = c \cdot Me_i$ for some constant vector $c = (c_1, \dots, c_d)$. Then V is the unique viscosity solution to (2.5) in the class of $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$.*

Proof of Theorem 2.1(a) is done in Lemma 3.1, Lemma 3.2 and Theorem 3.1. These results are proved in Section 3. Let us mention again that the uniqueness of viscosity solution does not follow from existing literature. We construct an appropriate test function, as stated in Theorem 4.2, in Section 5, which leads us to the uniqueness result.

2.1. Motivation from queueing network. The motivation of the above mentioned problem comes from a control problem in multi-class queueing network. Consider a queueing system with d customer classes where the arrivals of the customers are given by d independent Poisson processes $\{E^1, \dots, E^d\}$. There are d servers and every customer class has service priority in one of the servers. Therefore we can label the servers from 1 to d so that customer of class i has priority for service in server i for $i \in \mathcal{I}$. A customer of class j , $j \neq i$, may access the server i if there are no customers of class i present in the system. We assume that the network is working under a preemptive scheduling, i.e., service of a non-priority class customer is preempted by a priority class customer. Such queueing systems might be referred to as *queueing networks with help*. Similar type queueing network is studied in [4] for a M/M/N+M queueing network with ergodic type costs. The service time distributions of the customers at different servers are assumed to be exponential. If we denote by X_t^i , the number of class i customers in the system at time t , then we have the following balance equation

$$X_t^i = X_0^i + E_t^i - S^i(\mu^i \int_0^t B_s^i ds) - \sum_{j: j \neq i} S^{ij}(\mu^{ij} \int_0^t B_s^{ij} dI_s^j), \quad i \in \mathcal{I}. \quad (2.6)$$

In (2.6), $\{S^i, S^{ij}, i, j \in \mathcal{I}\}$ are independent standard Poisson process, μ^{ij} (μ^i) denotes the service rate of the class i (class i) customers at the server $j, j \neq i$ (i , respectively). B_s^{ij} (B_s^i) is an adapted process which denotes the status of service of class i customers at server j (i , respectively) at time s . $B^{ij} = 0$ implies that class i customer is not receiving service from server j whereas $B^{ij} \in (0, 1]$ implies that the class i customer is being served at server j with fraction of effort B^{ij} . Therefore we have

$$0 \leq \sum_{i:i \neq j} B_s^{ij} \leq 1.$$

Due to priority it is easy to see that $B^i \in \{0, 1\}$. I^i denote the cumulative time when there were no customers of class i present in the system, i.e.,

$$I_t^i := \int_0^t (1 - B_s^i) ds, \quad i \in \mathcal{I}.$$

Therefore I^i is an increasing process and integrations in (2.6) make sense. We assume that every class is in heavy traffic in the sense that the arrival rate of class i customers is equal to μ^i . Then we would have $\frac{I_{nt}^i}{n} \rightarrow 0$ as $n \rightarrow \infty$, and $t \geq 0$. We define the diffusion scaling as follows.

$$\begin{aligned} \hat{E}_t^i &:= \frac{1}{\sqrt{n}}(E_{nt}^i - \mu^i nt), \quad \hat{S}^i(t) := \frac{1}{\sqrt{n}}(S^i(nt) - nt), \quad \hat{S}^{ij}(t) := \frac{1}{\sqrt{n}}(S^{ij}(nt) - nt), \\ \hat{X}_t^i &:= \frac{1}{\sqrt{n}}X_{nt}^i, \quad \hat{I}_t^i := \frac{1}{\sqrt{n}}I_{nt}^i, \quad \bar{I}_t^i := \frac{1}{n}I_{nt}^i. \end{aligned}$$

Hence using (2.6) and a simple calculation we obtain

$$\begin{aligned} \hat{X}_t^i &= \hat{X}_0^i + \hat{E}_t^i - \hat{S}^i(\mu^i \int_0^t B_{ns}^i ds) - \sum_{j:j \neq i} \hat{S}^{ij}(\mu^{ij} \int_0^t B_{ns}^{ij} d\bar{I}_s^j) \\ &\quad + \mu^i \hat{I}_t^i - \sum_{j:j \neq i} \mu^{ij} \int_0^t B_{ns}^{ij} d\hat{I}_s^j, \end{aligned} \tag{2.7}$$

for $i \in \mathcal{I}$. We also have

$$\int_0^t \hat{X}_s^i d\hat{I}_s^i = 0 \quad \text{for all } i \in \mathcal{I}, t \geq 0.$$

Now formally letting $n \rightarrow \infty$, in (2.7) we have: $(\hat{X}, \hat{I}) \rightarrow (Z, \hat{Z})$ and

$$\begin{aligned} Z_t^i &= Z_0^i + W_t^{i,1} - W_t^{i,2} + \mu^i \hat{Z}^i - \sum_{j:j \neq i} \mu^{ij} \int_0^t v_{ij}(s) d\hat{Z}_s^j, \\ \int_0^t Z_s^i d\hat{Z}_s^i &= 0, \quad \forall t \geq 0, \end{aligned} \tag{2.8}$$

where $\{W^{i,1}, W^{i,2}, i \geq 1\}$ are collection of independent Brownian motion with suitable variance. The relation between (2.7) and (2.8) is quite formal, in fact, the convergence

might not hold for any choice of control B_n^{ij} . However, this relation can be justified for a reasonable class of controls. Now if we redefine \hat{Z}^i as $\mu^i \hat{Z}^i$ then from (2.8) we get

$$\begin{aligned} Z_t^i &= Z_0^i + W^i + \hat{Z}^i - \sum_{j:j \neq i} \int_0^t \frac{\mu^{ij}}{\mu^j} v_{ij}(s) d\hat{Z}_s^j, \\ \int_0^\cdot Z_s^i d\hat{Z}_s^i &= 0, \quad \forall t \geq 0, \end{aligned} \tag{2.9}$$

for some Brownian motion (W^1, \dots, W^d) with suitable covariance matrix and v takes values in $\mathbf{M}(\mathbf{1})$. Assume that $\frac{\mu^{ij}}{\mu^j} < 1$ for all i, j . Then it is easy to see that (2.9) is a particular case of (2.1) and Condition 2.2 is also satisfied.

We can associate the following cost structure to the above queueing system.

$$V(\hat{X}_0) := \inf \mathbb{E} \left[\int_0^\infty e^{-\beta s} (\ell(\hat{X}_s) ds + \sum_{i \in \mathcal{I}} c_i d\hat{I}_s^i) \right],$$

for some non-negative constants $c_i, i \in \mathcal{I}$, where the infimum is taken over all adapted process B . We note that if we let $n \rightarrow \infty$ formally then the above cost structure is a particular form of (2.4). Cost structure of above type has been used to find optimal control for various queueing networks (see for example [1] and references therein). The importance of such control problems comes from well studied Generalized Processor Sharing (GPS) networks [12,13]. In a single server network working under GPS scheduling in heavy-traffic, the server spends a positive fraction of effort (corresponding to their traffic intensities) to serve each customer class when all classes have backlog of work, otherwise the service capacity is split among the existing classes in a certain proportion. It is easy to see that our model considers a full-fledged optimal control version of the GPS models.

Remark 2.2. The above relation between the queueing control problem and the value function defined in (2.4) is bit formal. We neither study the convergence of the value functions corresponding to the queueing model, nor try to find any asymptotic optimality result. These are topics of further study. The main goal of this article is to characterize the value function (2.4) in terms of the solution of HJB (2.5). However, in the many server settings above questions are addressed in [4]. The limiting HJB in [4] has a classical solution which is used to obtain an optimal control for the limiting problem. This optimal control is then used to obtain convergence result of the value functions for the queueing model and asymptotic optimality. It should be noted that the solution of (2.5) is obtained in viscosity sense which gives additional difficulty in establishing convergence results of value functions.

3. PROOF OF THEOREM 2.1(A)

Conditions 2.1 and 2.2 will be in full effect in the remaining part of the article. We start by proving the growth estimate of V given by (2.4).

Lemma 3.1. *Let V be given by (2.4). Then there exist $c_1 \in (0, \infty)$ such that*

$$|V(x)| \leq c_1(1 + |x|^m), \quad x \in \mathbb{R}_+^d.$$

Proof. Fix $x \in \mathbb{R}_+^d$ and $v \in \mathfrak{M}$. Let (X, Y) be the solution of (2.1) given the control v and initial data x . Then we have

$$X_i(t) = x_i + \int_0^t b_i(X(s)) ds + \sum_{j=1}^d \Sigma_{ij} W_j(t) + \int_0^t \sum_{j=1}^d v_{ij}(s) dY_j(s), \quad t \geq 0. \quad (3.1)$$

Therefore summing over i we have from (3.1) that

$$\sum_{i \in \mathcal{I}} X_i(t) = \sum_{i \in \mathcal{I}} x_i + \int_0^t \sum_{i \in \mathcal{I}} b_i(X(s)) ds + \sum_{j=1}^d \left(\sum_{i \in \mathcal{I}} \Sigma_{ij} \right) W_j(t) + \sum_{j=1}^d \left[\int_0^t \sum_{i \in \mathcal{I}} v_{ij}(s) dY_j(s) \right] \quad (3.2)$$

Since v takes values in \mathbf{M} we have $\sum_i v_{ij} \leq 1$, thus using Condition 2.1 and (3.2) we get a constant $\kappa_2 > 0$ such that

$$\sum_{i \in \mathcal{I}} X_i(t) \leq \sum_{i \in \mathcal{I}} x_i + \kappa_2 \left[t + \sum_{i \in \mathcal{I}} (|W_i(t)| + Y_i(t)) \right].$$

Now use (A.3) to obtain

$$\sum_{i \in \mathcal{I}} X_i(t) \leq \sum_{i \in \mathcal{I}} x_i + \kappa_3 \left(t + \sum_{i \in \mathcal{I}} \sup_{0 \leq s \leq t} |W_i(s)| \right), \quad (3.3)$$

for some constant κ_3 . Recall that for any $a_i \geq 0, i = 1, \dots, d+2$, and $m \in [0, \infty)$ we have

$$\left(\sum_{i=1}^{d+2} a_i \right)^m \leq (d+2)^{(m-1)^+} \sum_{i=1}^{d+2} a_i^m.$$

Since $X \geq 0$ we have from (2.2) and (3.3) that

$$\begin{aligned} \ell(X(t)) &\leq c_\ell + c_\ell \left[\sqrt{d}|x| + \kappa_3 \left(t + \sum_{i \in \mathcal{I}} \sup_{0 \leq s \leq t} |W_i(s)| \right) \right]^m \\ &\leq c_\ell + c_\ell (d+2)^{(m-1)^+} d^{m/2} |x|^m + c_\ell (d+2)^{(m-1)^+} \kappa_3 \left(t^m + \sum_{i \in \mathcal{I}} \sup_{0 \leq s \leq t} |W_i(s)|^m \right) \\ &\leq c_\ell (d+2)^{(m-1)^+} d^{m/2} (1 + |x|^m) + c_\ell (d+2)^{(m-1)^+} \kappa_3 \left(t^m + \sum_{i \in \mathcal{I}} \sup_{0 \leq s \leq t} |W_i(s)|^m \right). \end{aligned} \quad (3.4)$$

Now we can find constants $\kappa_i, i = 4, 5, 6$, depending only on m , so that

$$\begin{aligned} \mathbb{E}_x \left[\int_0^\infty e^{-\beta t} (t^m + \sum_{i \in \mathcal{I}} \sup_{0 \leq s \leq t} |W_i(s)|^m) dt \right] &\leq \kappa_4 \left(1 + \int_0^\infty e^{-\beta t} \mathbb{E} \left[\sup_{0 \leq s \leq t} |W_1(s)|^m \right] \right) ds \\ &\leq \kappa_5 \left(1 + \int_0^\infty e^{-\beta t} t^{m/2} \right) ds \\ &\leq \kappa_6, \end{aligned}$$

where in the second inequality we use Burkholder-Davis-Gundy inequality [10]. Therefore combining the above display with (3.4) we obtain

$$\mathbb{E} \left[\int_0^\infty e^{-\beta t} \ell(X(t)) dt \right] \leq \kappa_7 (1 + |x|^m), \quad (3.5)$$

for some positive constant κ_7 . Let $|h_i|_\infty := \sup_{M \in \mathbf{M}} |h_i(M)|$ and $|b|_1 = \sum_i \sup_{x \in \mathbb{R}_+^d} |b_i(x)|$. Then for each $i \in \{1, \dots, d\}$,

$$\begin{aligned}
\left| \mathbb{E}_x \left[\int_0^\infty e^{-\beta t} h_i(v(t)) dY_i(t) \right] \right| &\leq |h_i|_\infty \left| \mathbb{E}_x \left[\int_0^\infty e^{-\beta t} dY_i(t) \right] \right| \\
&= |h_i|_\infty \left| \mathbb{E}_x \left[\int_0^\infty \beta e^{-\beta t} Y_i(t) dt \right] \right| \\
&\leq \kappa_8 \frac{|h_i|_\infty}{1 - \max_i \alpha_i} \left| \mathbb{E} \left[\int_0^\infty \beta e^{-\beta t} (|b|_1 t + \sum_{i \in \mathcal{I}} \sup_{0 \leq s \leq t} |W_i(s)|) dt \right] \right| \\
&\leq \kappa_9,
\end{aligned} \tag{3.6}$$

for some constants κ_8, κ_9 , where the second equality is obtained by a use of an integration-by-parts formula (this is justified by (A.3) and Law of Iterated Logarithm which confirms that maximum of a standard Brownian motion has polynomial growth almost surely), in the third inequality we use (A.3), and last inequality can be obtained following similar estimate as above (display above (3.5)). We also note that the constants above do not depend on i and $v \in \mathfrak{M}$. Therefore we complete the proof combining (3.5) and (3.6). \square

Next lemma establishes continuity of the value function V .

Lemma 3.2. *The value function V given by (2.4) is continuous in \mathbb{R}_+^d . In particular, $V \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$.*

Proof. To show the continuity we prove the following: for any sequence $\{v^n\} \in \mathfrak{M}$ and $x_n \rightarrow x \in \mathbb{R}_+^d$, as $n \rightarrow \infty$, we have

$$|J(x_n, v^n) - J(x, v^n)| \rightarrow 0, \quad \text{as } n \rightarrow \infty, \tag{3.7}$$

where J is given by (2.3). From the proof of Lemma 3.1 it is clear that

$$\limsup_{T \rightarrow \infty} \sup_{v \in \mathfrak{M}} \mathbb{E}_x \left[\int_T^\infty e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v(s)) dY_i(s)) \right] = 0.$$

Therefore to prove (3.7) it is enough to show that for any $T > 0$,

$$\begin{aligned}
&\mathbb{E}_x \left[\int_0^T e^{-\beta s} (\ell(X^n(s)) ds + \sum_{i \in \mathcal{I}} h_i(v^n(s)) dY_i^n(s)) \right] \\
&\quad - \mathbb{E}_x \left[\int_0^T e^{-\beta s} (\ell(\tilde{X}^n(s)) ds + \sum_{i \in \mathcal{I}} h_i(v^n(s)) d\tilde{Y}_i^n(s)) \right] \rightarrow 0,
\end{aligned} \tag{3.8}$$

as $n \rightarrow \infty$ where $(X^n, Y^n), (\tilde{X}^n, \tilde{Y}^n)$ are the solutions of (2.1) with the control v^n and the initial data x_n, x , respectively. It should be noted that both the solutions $(X^n, Y^n), (\tilde{X}^n, \tilde{Y}^n)$ are governed by the control v^n for all n . Also from (3.4) and the calculations in (3.6) we see that dominated convergence theorem can be applied to establish (3.8) if we can show

the following hold,

$$\begin{aligned} & \left| \int_0^T e^{-\beta s} (\ell(X^n(s)) ds - \int_0^T e^{-\beta s} (\ell(\tilde{X}^n(s)) ds \right| \rightarrow 0 \quad a.s., \\ & \left| \int_0^T e^{-\beta s} \sum_{i \in \mathcal{I}} h_i(v^n(s)) dY_i^n(s) - \int_0^T \sum_{i \in \mathcal{I}} e^{-\beta s} h_i(v^n(s)) d\tilde{Y}_i^n(s) \right| \rightarrow 0 \quad a.s. \end{aligned} \quad (3.9)$$

We recall the maps $(\mathcal{T}, \mathcal{S})$ from Theorem A.1. We have

$$Y_i^n(t) = \mathcal{T}_i(X^n, Y^n)(t) = \sup_{0 \leq s \leq t} \max\{0, -U_i(X^n, Y^n)(s)\},$$

$$X_i^n(t) = \mathcal{S}_i(X^n, Y^n)(t) = U_i(X^n, Y^n)(t) + \mathcal{T}_i(X^n, Y^n)(t),$$

$$\text{where } U_i(X^n, Y^n)(t) = (x_n)_i + \int_0^t b_i(X^n(s)) ds + (\Sigma W)_i(t) + \sum_{j \neq i} \int_0^t v_{ij}^n(s) dY_j^n(s).$$

The above relation also holds if we replace $(X^n, Y^n), x_n$ by $(\tilde{X}^n, \tilde{Y}^n)$ and x , respectively. Let $[\cdot]_{st}$ denote the bounded variation norm on the interval $[s, t]$. Then using the result from [16, pp. 171] we have for all $i \in \mathcal{I}$,

$$\begin{aligned} [Y_i^n - \tilde{Y}_i^n]_{st} & \leq |Y_i^n(s) - \tilde{Y}_i^n(s)| + |U_i(X^n, Y^n)(s) - U_i(\tilde{X}^n, \tilde{Y}^n)(s)| \\ & \quad + [U_i(X^n, Y^n) - U_i(\tilde{X}^n, \tilde{Y}^n)]_{st}. \end{aligned}$$

Now one can follow similar calculation as in (A.4) to obtain

$$\begin{aligned} [Y^n - \tilde{Y}^n]_{st} & \leq \sum_{i \in \mathcal{I}} |Y_i^n(s) - \tilde{Y}_i^n(s)| + sKd\|X^n - \tilde{X}^n\|_s + Kd(t-s) \sup_{r \in [s, t]} \|X^n(r) - \tilde{X}^n(r)\| \\ & \quad + \max_i \alpha_i [Y^n - \tilde{Y}^n]_{st}, \end{aligned}$$

which gives

$$[Y^n - \tilde{Y}^n]_{st} \leq \frac{1}{1 - \max_i \alpha_i} \left(\sum_{i \in \mathcal{I}} |Y_i^n(s) - \tilde{Y}_i^n(s)| + Kds\|X^n - \tilde{X}^n\|_s + Kd(t-s) \sup_{r \in [s, t]} \|X^n(r) - \tilde{X}^n(r)\| \right). \quad (3.10)$$

Similarly, we get

$$\begin{aligned} \sup_{r \in [s, t]} \|X^n(r) - \tilde{X}^n(r)\| & \leq \|X^n(s) - \tilde{X}^n(s)\| + Kd(t-s) \sup_{r \in [s, t]} \|X^n(r) - \tilde{X}^n(r)\| \\ & \quad + (1 + \max_i \alpha_i) [Y^n - \tilde{Y}^n]_{st} + \sum_{i \in \mathcal{I}} |Y_i^n(s) - \tilde{Y}_i^n(s)| \end{aligned}$$

and using (3.10) we have

$$\begin{aligned} \sup_{r \in [s, t]} \|X^n(r) - \tilde{X}^n(r)\| & \leq \left(1 + \frac{2 \vee (2Kds)}{1 - \max_i \alpha_i}\right) (\|X^n(s) - \tilde{X}^n(s)\| + \sum_{i \in \mathcal{I}} |Y_i^n(s) - \tilde{Y}_i^n(s)|) \\ & \quad + Kd(t-s) \left(1 + \frac{2}{1 - \max_i \alpha_i}\right) \sup_{r \in [s, t]} \|X^n(r) - \tilde{X}^n(r)\|. \end{aligned} \quad (3.11)$$

Now we choose $\theta > 0$ so that $Kd\theta(1 + \frac{2}{1 - \max_i \alpha_i}) < 1$. Since $x_n \rightarrow x$, we obtain from (3.10) and (3.11) that $\|X^n - \tilde{X}^n\|_\theta, [Y^n - \tilde{Y}^n]_\theta$ tends to 0 as $n \rightarrow \infty$. Similar convergence can

be extended on interval $[0, k\theta]$, k is a positive integer, by an iterative procedure and using (3.10) and (3.11). This proves (3.9). \square

Now we prove the dynamic programming principle for V . Recall that $\mathbb{B}_r(x)$ denotes the ball of radius r around x in \mathbb{R}_+^d . Let (X, Y) be a solution of (2.1) with initial data x . Denote

$$\sigma_r := \inf\{t \geq 0 : X(t) \notin \mathbb{B}_r(x)\}. \quad (3.12)$$

That is, σ_r denotes the exit time of X from the ball $\mathbb{B}_r(x)$.

Proposition 3.1 (Dynamic programming principle). *Let $\sigma = \sigma_r$ be as above. Then for any $t > 0$ we have*

$$V(x) = \inf_{v \in \mathfrak{M}} \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v(s)) dY_i(s)) + e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) \right]. \quad (3.13)$$

Proof. Using the Markov property of Brownian motion one gets that for any $v \in \mathfrak{M}$,

$$\mathbb{E}_x \left[\int_{\sigma \wedge t}^{\infty} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \geq e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)). \quad (3.14)$$

See for example [17, Lemma 3.2]. For $\varepsilon > 0$, we have $v \in \mathfrak{M}$ such that

$$\begin{aligned} V(x) &\geq J(x, v) - \varepsilon \\ &= \mathbb{E} \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) + \right. \\ &\quad \left. \mathbb{E}_x \left[\int_{\sigma \wedge t}^{\infty} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \right] - \varepsilon \\ &\geq \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) + e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) \right] - \varepsilon, \end{aligned}$$

where in the last inequality we use (3.14). Since ε is arbitrary, we get

$$V(x) \geq \inf_{v \in \mathfrak{M}} \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) + e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) \right]. \quad (3.15)$$

To prove the other direction we use Lemma A.1 from Appendix which says that for given $\varepsilon > 0$, $v \in \mathfrak{M}$, there exist control $\tilde{v} \in \mathfrak{M}$ such that

$$\mathbb{E}_x \left[\int_{\sigma \wedge t}^{\infty} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(\tilde{v}) dY_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \leq e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) + 3\varepsilon,$$

and

$$v(s) = \tilde{v}(s) \quad \text{for all } s \leq \sigma \wedge t.$$

Then

$$\begin{aligned} V(x) &\leq \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) + \right. \\ &\quad \left. \mathbb{E}_x \left[\int_{\sigma \wedge t}^{\infty} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(\tilde{v}) dY_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \right] \end{aligned}$$

$$\leq \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v) dY_i) + e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) \right] + 3\varepsilon.$$

Now we take the infimum over v to get the other side inequality. Hence (3.13) follows using (3.15) and the above display. \square

Lemma 3.3. *Given $r, \varepsilon \in (0, 1)$ there exists $t \in (0, 1)$ such that*

$$\sup_{v \in \mathfrak{M}} \mathbb{P}_x(\sigma_r \leq t) < \varepsilon \quad \text{for all } x \in \mathbb{R}_+^d,$$

where σ_r is given by (3.12).

Proof. Fix $x \in \mathbb{R}^d$ and $v \in \mathfrak{M}$. Let (X, Y) be the solution of (2.1) given the control v and initial data x . Then for $i \in \mathcal{I}$ we have

$$X_i(t) = x_i + \int_0^t b_i(X(s)) ds + (\Sigma W(t))_i + \sum_{j=1}^d \int_0^t v_{ij}(s) dY_j(s), \quad t \geq 0. \quad (3.16)$$

Then summing over i we get from (3.16) that

$$\begin{aligned} \sum_{i \in \mathcal{I}} |X_i(t) - x_i| &\leq t \sum_{i \in \mathcal{I}} \sup_{x \in \mathbb{R}_+^d} |b_i(x)| + \sum_{i \in \mathcal{I}} |(\Sigma W(t))_i| + (1 + \max_i \alpha_i) \sum_{i \in \mathcal{I}} Y_j(t) \\ &\leq \kappa_1 \left(t \sup_{x \in \mathbb{R}_+^d} |b(x)| + \sum_{i \in \mathcal{I}} \sup_{s \leq t} |W_i(s)| \right), \end{aligned} \quad (3.17)$$

for some constant κ_1 , depending on Σ, α , where in the last inequality we use (A.3). Now if $|X(t) - x| \geq r$ and $t \kappa_1 \sup_{x \in \mathbb{R}_+^d} |b(x)| < r/2$ then using (3.17) we obtain

$$r/2 < \kappa_1 \sum_{i \in \mathcal{I}} \sup_{s \leq t} |W_i(s)|.$$

Thus we have for the above choice t that

$$\begin{aligned} \sup_{v \in \mathfrak{M}} \mathbb{P}_x \left(\sup_{s \leq t} |X(s) - x| \geq r \right) &\leq \mathbb{P} \left(\sum_{i \in \mathcal{I}} \sup_{s \leq t} |W_i(s)| > \frac{r}{2\kappa_1} \right) \\ &\leq \frac{4d^2 \kappa_1^2}{r^2} \mathbb{E} \left[\sup_{s \leq t} |W_1(s)|^2 \right] \\ &\leq \frac{16d^2 \kappa_1^2}{r^2} t, \end{aligned}$$

where in the last line we use Doob's martingale inequality. Thus we can further restrict t , depending on ε , so that the rhs of above display is smaller than ε . This completes the proof observing that $\mathbb{P}_x(\sigma_r \leq t) \leq \mathbb{P}_x(\sup_{s \leq t} |X(s) - x| \geq r)$. \square

Now we are ready to show that the value function V , given by (2.4), is a viscosity solution. This clearly concludes the proof of Theorem 2.1 (a).

Theorem 3.1. *Assume that Conditions 2.1–2.3 hold. The value function V is a viscosity solution of the HJB given by (2.5).*

Proof. First we show that V is a super-solution. Let $\varphi \in \mathcal{C}^2(\mathbb{R}_+^d)$ be such that $V - \varphi$ attains its local minimum at $x \in \mathbb{R}_+^d$. We can also assume that $V(x) = \varphi(x)$, otherwise we have to translate φ . Therefore we need to show that

$$(\mathcal{L}\varphi(x) - \beta\varphi(x) + \ell(x)) \wedge \min_{x \in \mathbf{I}(x)} \mathcal{H}_i(D\varphi(x)) \leq 0. \quad (3.18)$$

Let us assume that (3.18) does not hold and we have a $\theta > 0$ such that

$$(\mathcal{L}\varphi(x) - \beta\varphi(x) + \ell(x)) \wedge \min_{x \in \mathbf{I}(x)} \mathcal{H}_i(D\varphi(x)) \geq 2\theta.$$

Using the continuity property we can find $r > 0$ such that the following holds

$$\begin{aligned} \mathcal{L}\varphi(y) - \beta\varphi(y) + \ell(y) &\geq \theta \quad \text{for all } y \in \mathbb{B}_r(x), \quad V - \varphi \geq 0 \text{ on } \mathbb{B}_r(x), \\ \mathcal{H}_i(D\varphi(y)) &\geq \theta \quad \text{for } y \in \mathbb{B}_r(x), \text{ and } \mathbf{I}(y) \subset \mathbf{I}(x) \text{ for } y \in \mathbb{B}_r(x). \end{aligned} \quad (3.19)$$

From (3.19) we get

$$D\varphi(y) \cdot Me_i + h_i(M) \geq \theta \quad \text{for all } M \in \mathbf{M}, \quad i \in \mathbf{I}(y), \quad y \in \mathbb{B}_r(x). \quad (3.20)$$

Let $v \in \mathfrak{M}$ and (X, Y) be the corresponding solution to (2.1) with initial data x . Define $\sigma = \sigma_r = \inf\{t \geq 0 : X(t) \notin \mathbb{B}_r(x)\}$. Apply Itô formula [10] to obtain

$$\begin{aligned} V(x) = \varphi(x) &= \mathbb{E}_x[e^{-\beta\sigma \wedge t} \varphi(X(\sigma \wedge t))] \\ &\quad - \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\mathcal{L}\varphi(X(s)) - \beta\varphi(X(s)) + D\varphi(s) \cdot v(s) dY(s)) ds \right] \\ &\leq \mathbb{E}_x[e^{-\beta\sigma \wedge t} V(X(\sigma \wedge t))] \\ &\quad + \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) - \theta ds - D\varphi(s) \cdot v(s) dY(s)) ds \right], \end{aligned}$$

where in the last line we use (3.19). Again using (3.20) we see that

$$\begin{aligned} \int_0^{\sigma \wedge t} e^{-\beta s} D\varphi(X(s)) \cdot v(s) dY(s) &= \sum_{i \in \mathbf{I}} \int_0^{\sigma \wedge t} e^{-\beta s} D\varphi(X(s)) \cdot v(s) e_i dY_i(s) \\ &\geq \sum_{i \in \mathbf{I}} \int_0^{\sigma \wedge t} e^{-\beta s} (\theta - h_i(v(s))) dY_i(s), \end{aligned}$$

where we use the fact that $\int_0^{\sigma \wedge t} \sum_{i \notin \mathbf{I}(X(s))} dY_i(s) = 0$. Thus combining above two displays we obtain that for any $v \in \mathfrak{M}$,

$$\begin{aligned} V(x) &\leq \mathbb{E}_x \left[e^{-\beta\sigma \wedge t} V(X(\sigma \wedge t)) + \int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathbf{I}} h_i(v(s)) dY_i(s)) \right] \\ &\quad - \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\theta ds + \theta \sum_{i \in \mathbf{I}} dY_i(s)) \right]. \end{aligned}$$

Since $v \in \mathfrak{M}$ is arbitrary we have from above that

$$V(x) \leq \inf_{v \in \mathfrak{M}} \mathbb{E}_x \left[e^{-\beta\sigma \wedge t} V(X(\sigma \wedge t)) + \int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathbf{I}} h_i(v(s)) dY_i(s)) \right] - \alpha(t), \quad (3.21)$$

where

$$\alpha(t) = \inf_{v \in \mathfrak{M}} \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\theta ds + \theta \sum_{i \in \mathcal{I}} dY_i(s)) \right].$$

In view Lemma 3.3 we can find t so that $\alpha(t) > 0$ which will give us from (3.21)

$$V(x) < \inf_{v \in \mathfrak{M}} \mathbb{E}_x \left[e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) + \int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v(s)) dY_i(s)) \right],$$

but this is contradicting to Proposition 3.1. This establishes (3.18).

Now we show that V is also a subsolution. Let $\varphi \in \mathcal{C}^2(\mathbb{R}^d)$ be such that $V - \varphi$ attains a local maximum at the point $x \in \mathbb{R}_+^d$. Also assume that $V(x) = \varphi(x)$. We have to show that

$$(\mathcal{L}\varphi(x) - \beta\varphi(x) + \ell(x)) \vee \max_{i \in \mathbf{I}(x)} \mathcal{H}_i(D\varphi(x)) \geq 0. \quad (3.22)$$

If not, then we can find $\theta > 0$ such that

$$(\mathcal{L}\varphi(x) - \beta\varphi(x) + \ell(x)) \vee \max_{i \in \mathbf{I}(x)} \mathcal{H}_i(D\varphi(x)) \leq -2\theta. \quad (3.23)$$

By Condition 2.3 there exists $M \in \mathbf{M}$ such that

$$\max_{i \in \mathbf{I}(x)} \{D\varphi(x) \cdot Me_i + h_i(M)\} < -\theta.$$

Therefore using (3.23) we can find $r > 0$ such that $V(y) - \varphi(y) \leq 0$ for $y \in \mathbb{B}_r(x)$ and the following hold:

$$\begin{aligned} \mathcal{L}\varphi(y) - \beta\varphi(y) + \ell(y) &< -\theta/2 \quad \text{for } y \in \mathbb{B}_r(x), \\ \{D\varphi(y) \cdot Me_i + h_i(M)\} &< -\theta/2 \quad \text{for } i \in \mathbf{I}(y), \quad \mathbf{I}(y) \subset \mathbf{I}(x), \text{ and } y \in \mathbb{B}_r(x). \end{aligned} \quad (3.24)$$

Now we take the control $v \equiv M$ and follow the same calculations as above, and using (3.24), to arrive at

$$\begin{aligned} V(x) &\geq \mathbb{E}_x \left[e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) + \int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v(s)) dY_i(s)) \right] \\ &\quad + \mathbb{E}_x \left[\int_0^{\sigma \wedge t} e^{-\beta s} (\theta/2 ds + \theta/2 \sum_{i \in \mathcal{I}} dY_i(s)) \right]. \end{aligned}$$

Now using Lemma 3.3 we know that the last term on rhs of above display is strictly positive for a suitably chosen t and therefore we obtain

$$V(x) > \mathbb{E}_x \left[e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) + \int_0^{\sigma \wedge t} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(v(s)) dY_i(s)) \right],$$

but this is contradicting to Proposition 3.1. This proves (3.22). Hence V is a viscosity solution to (2.5). \square

4. PROOF OF THEOREM 2.1(B)

In this section we show the uniqueness property of the viscosity solutions of (2.5). Indeed, we prove the following theorem.

Theorem 4.1. *Let Conditions 2.1 and 2.2 hold. Assume that there exists a vector $c = (c_1, \dots, c_d) \in \mathbb{R}^d$ such that $h_i(M) = c \cdot Me_i$ for all $i \in \mathcal{I}$. If $u \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ and $v \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ are, respectively, a viscosity subsolution and a viscosity supersolution to (2.5), then $u \leq v$ on \mathbb{R}_+^d .*

Theorem 2.1 (b) follows immediately from the theorem above.

Proof of Theorem 2.1 (b). According to Theorem 2.1 (a) (or Theorem 3.1) and Lemma 3.1 assure that the value function V given by (2.4) is a viscosity solution to (2.5) in the class $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$. Theorem 4.1 then guarantees that V is the unique viscosity solution to (2.5) in $\mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$. \square

To establish the comparison theorem above, it is crucial to find a test function having a few of properties, stated in the next theorem.

Theorem 4.2. *Assume that $h_i(M) \equiv 0$ for all $i \in \mathcal{I}$. Then there exists a globally $\mathcal{C}^{1,1}$ function φ on \mathbb{R}^d having the properties:*

$$\begin{aligned} \varphi(y) &> 0 \quad \text{for all } y \in \mathbb{R}^d \setminus \{0\}, \\ \varphi(\lambda y) &= \lambda^2 \varphi(y) \quad \text{for all } y \in \mathbb{R}^d \text{ and } \lambda \geq 0, \end{aligned}$$

and for each $i \in \mathcal{I}$,

$$\mathcal{H}_i(D\varphi(x)) \begin{cases} \geq 0 & \text{if } x_i \geq 0, \\ \leq 0 & \text{if } x_i \leq 0. \end{cases}$$

We note that if $d > 1$ and $h_i = 0$ for some $i \in \mathcal{I}$, then

$$\mathcal{H}_i(p) = \min_{M \in \mathbf{M}(\alpha)} p \cdot Me_i = p_i - \alpha_i \max_{j \in \mathcal{I} \setminus \{i\}} p_j^+,$$

while if $d = 1$ and $h_1 = 0$, then

$$\mathcal{H}_1(p) = p.$$

The proof of the theorem above is given in the next section.

Proof of Theorem 4.1. Let $u \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ and $v \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$ be, respectively, a viscosity subsolution and a viscosity supersolution to (2.5).

First of all, we note that it can be assumed that $c = 0$ and $\beta = 1$. Indeed, setting $u_1(x) = u(x) + c \cdot x$ and $v_1(x) = v(x) + c \cdot x$, observing that $w := u_1$ (resp., $w := v_1$) is a viscosity subsolution (resp., supersolution) to the HJB

$$\begin{cases} \mathcal{L}w - \beta w + \ell(x) + c \cdot (-b(x) + \beta x) = 0 & \text{in } (\mathbb{R}_+^d)^\circ, \\ \min_{M \in \mathbf{M}(\alpha)} Dw \cdot Me_i = 0 & \text{for } x \in \partial \mathbb{R}_+^d, i \in \mathbf{I}(x). \end{cases}$$

This boundary value problem can be rewritten as

$$\begin{cases} \beta^{-1}\mathcal{L}w - w + \beta^{-1}(\ell(x) - c \cdot b(x)) + c \cdot x = 0 & \text{in } (\mathbb{R}_+^d)^\circ, \\ \min_{M \in \mathbf{M}(\alpha)} Dw \cdot Me_i = 0 & \text{for } x \in \partial\mathbb{R}_+^d, \ i \in \mathbf{I}(x). \end{cases}$$

Thus, replacing u, v, A, b and ℓ by $u_1, v_1, \beta^{-1}A, \beta^{-1}b$ and $\beta^{-1}(\ell - c \cdot b) + c \cdot x$, if necessary, we may assume that $c = 0$ and $\beta = 1$.

Henceforth we assume that $c = 0$ and $\beta = 1$. In particular, we have $h_i = 0$ for all $i \in \mathcal{I}$. Since $u, v \in \mathcal{C}_{\text{pol}}(\mathbb{R}_+^d)$, there are constants $m \in \mathbb{N}$ and $\kappa > 0$ such that

$$|u(x)| \vee |v(x)| \leq \kappa(1 + |x|)^m \quad \text{for all } x \in \mathbb{R}_+^d.$$

Let $\varepsilon > 0$ and set for $x \in \mathbb{R}_+^d$,

$$\begin{aligned} \psi(x) &= (m+1)^{-1}(x_1^{m+1} + \cdots + x_d^{m+1}) + \mathbf{1} \cdot x, \\ u_\varepsilon(x) &= u(x) - \varepsilon(C + \psi(x)) \quad \text{and} \quad v_\varepsilon(x) = v(x) + \varepsilon(C + \psi(x)), \end{aligned}$$

where C is a constant to be fixed soon.

Note that $\psi \in \mathcal{C}^2(\mathbb{R}_+^d)$,

$$\lim_{|x| \rightarrow \infty} u_\varepsilon(x) = -\infty \quad \text{and} \quad \lim_{|x| \rightarrow \infty} v_\varepsilon(x) = \infty, \quad (4.1)$$

and observe that for some constants $C_i > 0$, with $i = 1, 2, 3$, and for all $x \in \mathbb{R}_+^d$,

$$|D\psi(x)| \leq C_1|x|^m, \quad |\mathcal{L}\psi(x)| \leq C_2(1 + |x|)^m \quad \text{and} \quad |\mathcal{L}\psi(x)| \leq C_3 + \psi(x),$$

where we have used the fact that b is bounded on \mathbb{R}_+^d .

Now, we fix $C = C_3$. Let $x \in \mathbb{R}_+^d$ and $(p, X) \in J^{2,+}u_\varepsilon(x)$. See [6] for the definition of semi-jets $J^{2,\pm}$. Setting $(q, Y) := (p, X) + \varepsilon(D\psi(x), D^2\psi(x)) \in J^{2,+}u(x)$, we compute that

$$\begin{aligned} \text{Tr } AX + b(x) \cdot p - u_\varepsilon(x) + \ell(x) & \\ &\geq \text{Tr } AY + b(x) \cdot q - u(x) + \ell(x) - \varepsilon|\mathcal{L}\psi(x)| + \varepsilon(C + \psi(x)) \\ &\geq \text{Tr } AY + b(x) \cdot q - u(x) + \ell(x). \end{aligned} \quad (4.2)$$

Recalling that

$$\mathcal{H}_i(p) = p_i - \alpha_i \max_{j \neq i} p_j^+ \quad \text{for all } i \in \mathcal{I},$$

and noting that for any $i \in \mathcal{I}$,

$$D_i\psi(x) = x_i^m - 1 \quad \text{for all } x \in \mathbb{R}_+^d,$$

where $D_i = \partial/\partial x_i$, we see that if $x_i = 0$, then

$$\begin{aligned} \mathcal{H}_i(p) &= q_i + \varepsilon - \alpha_i \max_{j \neq i} (q_j + \varepsilon(-x_j^m + 1))^+ \\ &\geq q_i + \varepsilon(1 - \alpha_i) - \alpha_i \max_{j \neq i} q_j^+ = \mathcal{H}_i(q) + \varepsilon(1 - \alpha_i). \end{aligned}$$

We thus deduce that $w := u_\varepsilon$ is a viscosity subsolution to

$$\begin{cases} \mathcal{L}w - w + \ell = 0 & \text{in } (\mathbb{R}_+^d)^\circ, \\ \mathcal{H}_i^-(Dw(x)) = 0 & \text{if } x \in \partial\mathbb{R}_+^d, \ i \in \mathbf{I}(x), \end{cases} \quad (4.3)$$

where $\mathcal{H}_i^-(p) := \mathcal{H}_i(p) - \varepsilon(1 - \alpha_i)$.

Similarly, we see that $w := v_\varepsilon$ is a viscosity supersolution to

$$\begin{cases} \mathcal{L}w - w + \ell = 0 & \text{in } (\mathbb{R}_+^d)^\circ, \\ \mathcal{H}_i^+(Dw(x)) = 0 & \text{if } x \in \partial\mathbb{R}_+^d, i \in \mathbf{I}(x), \end{cases} \quad (4.4)$$

where $\mathcal{H}_i^+(p) := \mathcal{H}_i(p) + \varepsilon(1 - \alpha_i)$.

It is enough to show that for any $\varepsilon > 0$, we have $u_\varepsilon \leq v_\varepsilon$ on \mathbb{R}_+^d .

To do this, we fix $\varepsilon > 0$ and suppose to the contrary that $\sup(u_\varepsilon - v_\varepsilon) > 0$. Let $\varphi \in \mathcal{C}^{1,1}(\mathbb{R}^d)$ be the test function given by Theorem 4.2. For any $k \in \mathbb{N}$, consider the function

$$\Phi(x, y) = u_\varepsilon(x) - v_\varepsilon(y) - k\varphi(x - y) \quad \text{on } (x, y) \in (\mathbb{R}_+^d)^2,$$

and let (x_k, y_k) be a maximum point of this function Φ . Note that the existence of such a maximum point is ensured by (4.1).

Since $\varphi \geq 0$ is globally $\mathcal{C}^{1,1}$ on \mathbb{R}^d , we find that for all $(x, y) \in (\mathbb{R}_+^d)^2$ and some constant $C > 0$,

$$\varphi(x - y) \leq \varphi(x_k - y_k) + D\varphi(x_k - y_k) \cdot (x - y - (x_k - y_k)) + C|x - y - (x_k - y_k)|^2.$$

It is standard to see (see [6] for instance) that

$$\lim_{k \rightarrow \infty} k\varphi(x_k - y_k) = 0 \quad (4.5)$$

and that there are symmetric matrices $X_k, Y_k \in \mathbb{R}^d \otimes \mathbb{R}^d$ such that

$$\begin{pmatrix} X_k & 0 \\ 0 & -Y_k \end{pmatrix} \leq 6Ck \begin{pmatrix} \mathbb{I} & -\mathbb{I} \\ -\mathbb{I} & \mathbb{I} \end{pmatrix}, \quad (4.6)$$

and

$$(p_k, X_k) \in \bar{J}^{2,+}u_\varepsilon(x_k) \quad \text{and} \quad (p_k, Y_k) \in \bar{J}^{2,-}v_\varepsilon(y_k), \quad (4.7)$$

where $p_k = kD\varphi(x_k - y_k)$ and $\bar{J}^{2,\pm}$ denotes the “closure” (see [6]) of the semi-jets $J^{2,\pm}$. The inequality above implies that $X_k \leq Y_k$.

Let $x_k = (x_{k1}, \dots, x_{kd})$ and $y_k = (y_{k1}, \dots, y_{kd})$. If $x_{ki} = 0$ for some $i \in \mathcal{I}$, then $x_{ki} - y_{ki} \leq 0$ and hence

$$\mathcal{H}_i^-(p_k) = k\mathcal{H}_i(D\varphi(x_k - y_k)) - \varepsilon(1 - \alpha_i) \leq -\varepsilon(1 - \alpha_i) < 0.$$

Thus, by the viscosity property of u_ε , we see that

$$\text{Tr } AX_k + b(x_k) \cdot p_k - u_\varepsilon(x_k) + \ell(x_k) \geq 0.$$

Similarly, we obtain

$$\text{Tr } AY_k + b(y_k) \cdot p_k - v_\varepsilon(y_k) + \ell(y_k) \leq 0.$$

Combining these two, we get

$$\begin{aligned} u_\varepsilon(x_k) - v_\varepsilon(y_k) &\leq \text{Tr } A(X_k - Y_k) + (b(x_k) - b(y_k)) \cdot p_k + \ell(x_k) - \ell(y_k) \\ &\leq K|x_k - y_k||p_k| + \ell(x_k) - \ell(y_k), \end{aligned} \quad (4.8)$$

where K is a Lipschitz bound of the function b . Furthermore, observe that $\{x_k, y_k\} \subset (\mathbb{R}_+^d)^2$ is a bounded sequence by (4.1) and in view of the homogeneity of φ that for all $z \in \mathbb{R}^d$ and for some constant $C_4 > 0$,

$$|z||D\varphi(z)| \leq C_4\varphi(z),$$

and, consequently,

$$|x_k - y_k||p_k| \leq C_4 k \varphi(x_k - y_k).$$

Thus, from (4.8) and (4.5), we infer that

$$\begin{aligned} 0 &< \sup_{\mathbb{R}_+^d} (u_\varepsilon - v_\varepsilon) \leq u_\varepsilon(x_k) - v_\varepsilon(y_k) - k\varphi(x_k - y_k) \\ &\leq K|x_k - y_k||p_k| + \ell(x_k) - \ell(y_k) \rightarrow 0 \quad \text{as } k \rightarrow \infty, \end{aligned}$$

which is a contradiction, and conclude that $u_\varepsilon \leq v_\varepsilon$ on \mathbb{R}_+^d . \square

Remark 4.1. The above proof also works if we replace the operator $\mathcal{L}V - \beta V + \ell$ by a general nonlinear operator satisfying some standard condition. One may look at [6, Conditions (0.1)-(0.3)], [7, Condition (1.3)], [8, Condition (1.4)] for general form of such operators.

5. CONSTRUCTION OF A TEST FUNCTION

We now give the proof of Theorem 4.2. That is, we construct a test function in the class $\mathcal{C}^{1,1}(\mathbb{R}^d)$, which is a main step in establishing the comparison principle.

When $d = 1$, the function $\varphi(x) = x^2$ obviously has the properties of test function stated in Theorem 4.2. We may thus assume in this section that $d > 1$.

5.1. A convex body. We set

$$\Xi^d := \{\xi = (\xi_1, \dots, \xi_d) : \xi_i = 1 \text{ or } \alpha_i \text{ for all } i \in \mathcal{I}\}.$$

For $\xi = (\xi_1, \dots, \xi_d) \in \Xi^d$, we set

$$\xi^\# = \#\{i \in \mathcal{I} : \xi_i = 1\},$$

where $\#A$ indicates the cardinality of the set A , and, for $k \in \{0, 1, \dots, d\}$,

$$\Xi_k^d := \{\xi \in \Xi^d : \xi^\# = k\}.$$

We set

$$\Xi_*^d := \bigcup_{k \in \mathcal{I}} \Xi_k^d = \Xi^d \setminus \{(\alpha_1, \dots, \alpha_d)\},$$

and note

$$\#\Xi_k^d = \binom{d}{k}, \quad \#\Xi_*^d = 2^d - 1.$$

Let

$$1 = \theta_1 < \dots < \theta_d,$$

and assume that

$$\theta_1 > \frac{d-1 + \max_{i \in \mathcal{I}} \alpha_i}{d} \theta_d. \tag{5.1}$$

Note that using Condition 2.2 we have

$$\theta_1 > \frac{d-1 + \max_{i \in \mathcal{I}} \alpha_i}{d} \theta_d > \frac{d \max_{i \in \mathcal{I}} \alpha_i}{d} \theta_d = \max_{i \in \mathcal{I}} \alpha_i \theta_d. \tag{5.2}$$

For $\xi \in \Xi_k^d$, we define

$$H(\xi) := \{x \in \mathbb{R}^d : \xi \cdot x \leq \theta_k\},$$

and, for $\delta > 0$, set

$$H_\delta^- := \{x \in \mathbb{R}^n : x_i \geq -\delta \text{ for all } 1 \leq i \leq d\},$$

$$\mathcal{S} := \bigcap_{\xi \in \Xi_*^d} H(\xi),$$

and

$$\mathcal{S}_\delta := \mathcal{S} \cap H_\delta^-.$$

Note that

$$\partial H(\xi) = \{x \in \mathbb{R}^d : \xi \cdot x = \theta_k\} \text{ for } \xi \in \Xi_k^d.$$

Lemma 5.1. *The set \mathcal{S}_δ is a compact convex subset of \mathbb{R}^d and is a neighborhood of the origin.*

Proof. It is clear that \mathcal{S}_δ is a closed convex subset of \mathbb{R}^d . Observe that

$$\mathcal{S}_\delta \subset H_\delta^- \cap H(\mathbf{1}),$$

and $H_\delta^- \cap H(\mathbf{1})$ is bounded in \mathbb{R}^d . Hence, \mathcal{S}_δ is a compact subset of \mathbb{R}^d .

The distance from the origin to the hyperplane

$$\partial H(\xi) = \{x \in \mathbb{R}^d : \xi \cdot x = \theta_k\},$$

with $\xi \in \Xi_k^d$, is given by

$$\theta_k/|\xi|.$$

Hence, setting

$$\rho := \delta \wedge \min\{\theta_k/|\xi| : \xi \in \Xi_k^d, k \in \mathcal{I}\} (\leq \delta),$$

we find that

$$B_\rho(0) \subset \mathcal{S}_\delta.$$

Thus, \mathcal{S}_δ is a neighborhood of the origin. □

Let $z \in \partial \mathcal{S}_\delta$ and let $N(z, \mathcal{S}_\delta)$ denote the normal cone of \mathcal{S}_δ at z , that is,

$$N(z, \mathcal{S}_\delta) := \{p \in \mathbb{R}^d : p \cdot (x - z) \leq 0 \text{ for all } x \in \mathcal{S}_\delta\}. \quad (5.3)$$

Set

$$J_z := \{i \in \mathcal{I} : z_i = -\delta\},$$

and

$$\Xi_z := \{\xi \in \Xi_*^d : z \in \partial H(\xi)\}.$$

According to [15, Corollary 23.8.1], we have

Lemma 5.2. *We have*

$$N(z, \mathcal{S}_\delta) = \text{conical hull}(\Xi_z \cup \{-e_j : j \in J_z\}) = \sum_{j \in J_z} \mathbb{R}_+(-e_j) + \mathbb{R}_+ \text{co} \Xi_z,$$

where $\text{co} \Xi_z$ denotes the convex hull of Ξ_z .

For the time being we examine the normal cone $N(z, \mathcal{S})$ at z for $z \in \partial\mathcal{S} \cap \partial\mathbb{R}_+^d$. Note that

$$\partial\mathbb{R}_+^d = \bigcup_{i \in \mathcal{I}} F_i,$$

where $F_i := \{x \in \mathbb{R}_+^d : x_i = 0\}$, and therefore

$$N(z, \mathcal{S}) = \mathbb{R}_+ \text{ co } \Xi_z.$$

Lemma 5.3. *Assume that $z \in F_i \cap \partial\mathcal{S}$. For any $\xi \in \text{co } \Xi_z$, we have*

$$\xi_i = \alpha_i \quad \text{and} \quad \max_{j \in \mathcal{I}} \xi_j = 1.$$

Proof. For notational convenience, we treat only the case when $i = d$. Let $\xi \in \text{co } \Xi_z$. In order to show that $\xi_d = \alpha_d$, we suppose to the contrary that $\xi_d \neq \alpha_d$. This implies that there exists $\bar{\xi} \in \Xi_z$ such that $\bar{\xi}_d = 1$. We select $k \in \mathcal{I}$ so that $\bar{\xi} \in \Xi_k^d$. Set

$$\hat{\xi} = (\bar{\xi}_1, \dots, \bar{\xi}_{d-1}, \alpha_d).$$

First we consider the case when $k > 1$. Note that $\hat{\xi} \in \Xi_{k-1}^d$ and that

$$\hat{\xi} \cdot z \leq \theta_{k-1} \quad \text{and} \quad \bar{\xi} \cdot z = \theta_k.$$

The last two relations are contradicting, since $\theta_{k-1} < \theta_k$ and

$$\hat{\xi} \cdot z = \hat{\xi} \cdot (z_1, \dots, z_{n-1}, 0) = \bar{\xi} \cdot (z_1, \dots, z_{n-1}, 0) = \bar{\xi} \cdot z.$$

Next consider the case when $k = 1$. Note that

$$\hat{\xi} = (\alpha_1, \dots, \alpha_d)$$

and

$$\theta_1 = \bar{\xi} \cdot z = \hat{\xi} \cdot z \leq \max_{i \in \mathcal{I}} \alpha_i \mathbf{1} \cdot z \leq \max_{i \in \mathcal{I}} \alpha_i \theta_d,$$

which contradicts (5.2).

Next we show that

$$\max_{i \in \mathcal{I}} \xi_i = 1.$$

We argue by contradiction, and suppose that

$$\max_{i \in \mathcal{I}} \xi_i < 1.$$

Since $\xi \in \text{co } \Xi_z$, there exist $\xi^1, \dots, \xi^N \in \Xi_z$ and $\lambda_i \geq 0$, with

$$\sum_{i=1}^N \lambda_i = 1 \quad \text{such that} \quad \xi = \sum_{i=1}^N \lambda_i \xi^i,$$

where we may assume that $N \geq d$.

Writing

$$\xi^i = (\xi_1^i, \dots, \xi_d^i) \quad \text{for } i \in \{1, \dots, N\},$$

by the supposition that

$$\max_{i \in \mathcal{I}} \xi_i < 1,$$

we find that for each $i \in \mathcal{I}$, there exists $j_i \in \{1, \dots, N\}$ such that

$$\xi_i^{j_i} = \alpha_i.$$

We can rearrange $(\xi^1, \lambda_1), \dots, (\xi^N, \lambda_N)$ in such a way that $N \geq d$ and $j_i \leq d$ for all $i \in \mathcal{I}$. For $i \in \mathcal{I}$, let $k_i \in \mathcal{I}$ be such that $\xi^i \in \Xi_{k_i}^d$. Observe that for any $i \in \mathcal{I}$,

$$\xi^i \cdot z = \theta_{k_i},$$

and hence

$$\frac{1}{d} \sum_{i \in \mathcal{I}} \xi^i \cdot z = \frac{1}{d} \sum_{i \in \mathcal{I}} \theta_{k_i} \geq \theta_1.$$

Since $z \in \mathcal{S} \cap F_d \subset H(1) \cap F_d$, we have

$$\frac{1}{d} \sum_{i \in \mathcal{I}} \xi^i \cdot z \leq \frac{(d-1) + \max \alpha_i}{d} \mathbf{1} \cdot z \leq \frac{(d-1) + \max \alpha_i}{d} \theta_d.$$

Thus combining above two displays, we get

$$\theta_1 \leq \frac{(d-1) + \max \alpha_i}{d} \theta_d.$$

This contradicts (5.1). □

Lemma 5.4. *The multi-valued map $z \mapsto \Xi_z$ is upper semicontinuous on $\partial \mathcal{S}$. That is, for each $z \in \partial \mathcal{S}$ there exists $\varepsilon > 0$ such that*

$$\Xi_x \subset \Xi_z \quad \text{for all } x \in B_\varepsilon(z) \cap \partial \mathcal{S}.$$

Proof. Take $z \in \partial \mathcal{S}$ and put

$$\Xi_z^c := \Xi_* \setminus \Xi_z.$$

Since

$$z \notin \partial H(\xi) \quad \text{for any } \xi \in \Xi_z^c$$

and $\#\Xi_z^c < \infty$, there is $\varepsilon > 0$ such that

$$\text{dist}(z, \partial H(\xi)) > \varepsilon \quad \text{for all } \xi \in \Xi_z^c.$$

That is, we have

$$B_\varepsilon(z) \cap \partial H(\xi) = \emptyset \quad \text{for all } \xi \in \Xi_z^c,$$

and, consequently,

$$\Xi_x \subset \Xi_z \quad \text{for all } x \in B_\varepsilon(z) \cap \partial \mathcal{S}.$$

□

Lemma 5.5. *There exists $\delta > 0$ such that for each $i \in \mathcal{I}$,*

$$\xi_i = \alpha_i \quad \text{and} \quad \max_{j \in \mathcal{I}} \xi_j = 1 \quad \text{for all } \xi \in \text{co } \Xi_z \quad \text{and} \quad z \in \partial \mathcal{S} \cap F_{i,\delta},$$

where $F_{i,\delta} = \{x \in \mathbb{R}^d : x_j \geq -\delta \text{ for all } j \in \mathcal{I}, |x_i| \leq \delta\}$.

Proof. Let $z \in \partial\mathcal{S} \cap F_d$. By Lemma 5.4, there exists $\varepsilon := \varepsilon_z > 0$ such that

$$\Xi_x \subset \Xi_z \quad \text{for all } x \in B_\varepsilon(z) \cap \partial\mathcal{S}.$$

This yields

$$\text{co } \Xi_x \subset \text{co } \Xi_z \quad \text{for all } x \in B_\varepsilon(z) \cap \partial\mathcal{S}.$$

The family $\{B_\varepsilon(z)\}$ is an open covering of the compact set $\partial\mathcal{S} \cap F_d$. Hence, there exists $\delta_d > 0$ such that for any $x \in \partial\mathcal{S} \cap F_{d,\delta_d}$, there is $z \in \partial\mathcal{S} \cap F_d$ such that

$$\text{co } \Xi_x \subset \text{co } \Xi_z,$$

which implies that

$$\xi_d = \alpha_d \quad \text{and} \quad \max_{i \in \mathcal{I}} \xi_i = 1 \quad \text{for all } \xi \in \text{co } \Xi_x.$$

Similarly, we may choose $\delta_i > 0$ for each $i \in \mathcal{I}$ such that for any $x \in \partial\mathcal{S} \cap F_{i,\delta_i}$,

$$\xi_i = \alpha_i \quad \text{and} \quad \max_{j \in \mathcal{I}} \xi_j = 1 \quad \text{for all } \xi \in \text{co } \Xi_x.$$

Thus, setting $\delta := \min_{i \in \mathcal{I}} \delta_i$, we conclude that for any $i \in \mathcal{I}$ and $x \in \partial\mathcal{S} \cap F_{i,\delta}$,

$$\xi_i = \alpha_i \quad \text{and} \quad \max_{j \in \mathcal{I}} \xi_j = 1 \quad \text{for all } \xi \in \text{co } \Xi_x.$$

□

5.2. Sign of \mathcal{H}_i . Let us recall that for $i \in \mathcal{I}$,

$$\mathcal{H}_i(p) = p_i - \alpha_i \max_{j \neq i} p_j^+.$$

Set

$$E_{i,\delta} = \{x \in \mathbb{R}^d : |x_i| \leq \delta\} \quad \text{for } i \in \mathcal{I}.$$

Lemma 5.6. *Let $\delta > 0$ be the constant from Lemma 5.5. Let $i \in \mathcal{I}$ and $x \in \partial\mathcal{S}_\delta \cap E_{i,\delta}$. Then*

$$\mathcal{H}_i(\gamma) \leq 0 \quad \text{for all } \gamma \in N(x, \mathcal{S}_\delta),$$

where $N(x, \mathcal{S}_\delta)$ is given by (5.3).

Proof. We set $J := \{j \in \mathcal{I} : x_j = -\delta\}$. Note that J or Ξ_x may be empty set. We note that

$$N(x, \mathcal{S}_\delta) = \sum_{j \in J} -\mathbb{R}_+ e_j + \mathbb{R}_+ \text{co } \Xi_x.$$

Let $\gamma \in N(x, \mathcal{S}_\delta)$. We choose $\lambda_j \geq 0$, $\lambda \geq 0$, and $\xi = (\xi_1, \dots, \xi_d) \in \text{co } \Xi_x$ so that

$$\gamma = \sum_{j \in J} -\lambda_j e_j + \lambda \xi.$$

By Lemma 5.5, we have

$$\xi_j = \alpha_j \quad \text{for all } j \in J,$$

and

$$\xi_i = \alpha_i \quad \text{and} \quad \max_{j \in \mathcal{I}} \xi_j = 1.$$

In particular, we have $\xi_k = 1$ for some $k \notin J \cup \{i\}$. Thus, we have

$$\gamma = \sum_{j \in J} (\lambda \alpha_j - \lambda_j) e_j + \lambda e_k + \sum_{l \in \mathcal{I} \setminus (J \cup \{k\})} \lambda \xi_l e_l.$$

Now, if $i \in J$, then

$$\mathcal{H}_i(\gamma) = \lambda \alpha_i - \lambda_i - \alpha_i \lambda = -\lambda_i \leq 0,$$

and, if $i \notin J$, then

$$\mathcal{H}_i(\gamma) = \lambda \alpha_i - \alpha_i \lambda = 0.$$

□

Set

$$G_i = \{x \in \mathbb{R}^d : x_i > -\delta\}.$$

Lemma 5.7. *Let $\delta > 0$ be as before. Let $i \in \mathcal{I}$ and $x \in \partial \mathcal{S}_\delta \cap G_i$. Then*

$$\mathcal{H}_i(\gamma) \geq 0 \quad \text{for all } \gamma \in N(x, \mathcal{S}_\delta).$$

Proof. Set $J := \{j \in \mathcal{I} : x_j = -\delta\}$. Note that J may be an empty set and that, since $x \in G_i$, we have $i \notin J$. Let $\gamma \in N(x, \mathcal{S}_\delta)$, and recall that

$$N(x, \mathcal{S}_\delta) = \sum_{j \in J} \mathbb{R}_+(-e_j) + \mathbb{R}_+ \text{co } \Xi_x.$$

Choose $\lambda_j \geq 0$, $\lambda \geq 0$, and $\xi \in \text{co } \Xi_x$ so that

$$\gamma = \sum_{j \in J} -\lambda_j e_j + \lambda \xi,$$

and observe that

$$\gamma = \sum_{j \in J} (\lambda \xi_j - \lambda_j) e_j + \sum_{k \in \mathcal{I} \setminus J} \lambda \xi_k e_k$$

and that

$$\alpha_k \leq \xi_k \leq 1 \quad \text{for all } k \in \mathcal{I}.$$

Hence,

$$\mathcal{H}_i(\gamma) = \gamma_i - \alpha_i \max_{k \neq i} \gamma_k^+ \geq \lambda \xi_i - \alpha_i \max_{k \neq i} \lambda \xi_k \geq \lambda \alpha_i - \alpha_i \lambda \max_{k \neq i} \xi_k \geq 0.$$

□

Remark 5.1. Combining Lemmas 5.6 and 5.7, we see that for any $x \in \partial \mathcal{S}_\delta$, if $|x_i| < \delta$ for some i , then

$$\mathcal{H}_i(\gamma) = 0 \quad \text{for all } \gamma \in N(x, \mathcal{S}_\delta).$$

5.3. Construction of $\varphi(x)$. For $0 < \varepsilon < \delta$, let

$$\mathcal{S}_\delta^\varepsilon := \left\{ x \in \mathbb{R}^d : \text{dist}(x, \mathcal{S}_\delta) \leq \varepsilon \right\},$$

where δ is the constant from Lemma 5.5. It is clear that $\mathcal{S}_\delta^\varepsilon$ satisfies the uniform interior sphere condition. Thus we have the following lemma.

Lemma 5.8. *The boundary $\partial\mathcal{S}_\delta^\varepsilon$ is locally represented as the graph of a $\mathcal{C}^{1,1}$ function.*

Proof. Fix any $z \in \partial\mathcal{S}_\delta^\varepsilon$. We show that $\partial\mathcal{S}_\delta^\varepsilon$ can be represented as the graph of a $\mathcal{C}^{1,1}$ function in a neighborhood of z . Making an orthogonal change of variables, we assume that $z = |z|e_d = (0, \dots, 0, |z|)$.

Let $B'_R(x')$ denote the open ball in \mathbb{R}^{d-1} of radius $R > 0$ and center at $x' \in \mathbb{R}^{d-1}$ and let $B'_R := B'_R(0)$. We choose $r > 0$ so that $\bar{B}'_{3r} \times \{0\} \subset (\mathcal{S}_\delta^\varepsilon)^\circ$. Recall that $\mathcal{S}_\delta^\varepsilon$ is a convex, compact, neighborhood of the origin of \mathbb{R}^d , note that, if $x' \in B'_{3r}$, then the set $\{x_d : (x', x_d) \in \mathcal{S}_\delta^\varepsilon\}$ is a closed, finite interval containing the origin of \mathbb{R} , and set

$$g(x') = \max\{x_d : (x', x_d) \in \mathcal{S}_\delta^\varepsilon\}.$$

Note that $0 < g(x') \leq M$ for all $x' \in B'_{3r}$ and some constant $M > 0$.

We show that $g \in \mathcal{C}^{1,1}(B'_r)$. Since $\mathcal{S}_\delta^\varepsilon$ is convex, g is concave on B'_{3r} . This together with the boundedness of g implies that g is Lipschitz continuous on B'_{2r} . Let $L > 0$ be a Lipschitz bound of g on B'_{2r} .

Fix any $x' \in B'_r$. By the definition of $\mathcal{S}_\delta^\varepsilon$ and the convexity of \mathcal{S}_δ , there exists a unique $y \in \mathcal{S}_\delta$ such that

$$\bar{B}_\varepsilon(x', g(x')) \cap \mathcal{S}_\delta = \{y\}.$$

On the other hand, since $y \in \mathcal{S}_\delta$, we have

$$B_\varepsilon(y) \subset \mathcal{S}_\delta^\varepsilon. \quad (5.4)$$

Now, using a simple geometry and the Lipschitz property of g , we deduce that

$$g(x') > y_d \quad \text{and} \quad \frac{|x' - y'|}{g(x') - y_d} \leq L.$$

Indeed, the second estimate above can be obtained as follows: due to the continuity of $x' \mapsto (x', g(x')) \mapsto y$, it is enough to establish the above relation at point x' where ∇g exists. Convexity of \mathcal{S}_δ and $\mathcal{S}_\delta^\varepsilon$ implies that $(x' - y', g(x') - y_d) \in N((x', g(x')), \mathcal{S}_\delta^\varepsilon)$ and $N((x', g(x')), \mathcal{S}_\delta^\varepsilon) = \{\lambda(-\nabla g(x'), 1) : \lambda \in \mathbb{R}_+\}$. Thus for some $\lambda > 0$ we have $(x' - y', g(x') - y_d) = \lambda(-\nabla g(x'), 1)$. Now it is easy to see that the second estimate above holds. Furthermore, we get

$$\varepsilon^2 = |x' - y'|^2 + (g(x') - y_d)^2 \geq (1 + L^{-2})|x' - y'|^2.$$

Thus, setting $\theta := L/\sqrt{1 + L^2} \in (0, 1)$, we have

$$|x' - y'| \leq \theta\varepsilon, \quad (5.5)$$

Note that $V := B'_r \cap B'_\varepsilon(y')$ is a neighborhood of x' . Let $z' \in V$ and observe by (5.4) that

$$|(z', g(z')) - y| \geq \varepsilon = |(x', g(x')) - y|.$$

From this, noting that $g(z') > y_d$, we get

$$g(z') \geq g(x') + \sqrt{\varepsilon^2 - |z' - y'|^2} - \sqrt{\varepsilon^2 - |x' - y'|^2}. \quad (5.6)$$

Next setting $h(z') = \sqrt{\varepsilon^2 - |z' - y'|^2}$ for $z' \in V$ and using (5.5), we calculate

$$D^2h(x') = -\frac{1}{\varepsilon^2 - |x' - y'|^2} \left(\mathbb{I}' + \frac{(x' - y') \otimes (x' - y')}{\sqrt{\varepsilon^2 - |x' - y'|^2}} \right),$$

where \mathbb{I}' denotes the $(d-1) \times (d-1)$ identity matrix, and moreover,

$$D^2h(x') \geq -\frac{1}{\varepsilon^2(1-\theta^2)} \left(1 + \frac{\varepsilon\theta^2}{\sqrt{1-\theta^2}} \right) \mathbb{I}'.$$

This and (5.6) show that g is semi-convex on B'_r , which together with the concavity of g ensures that $g \in \mathcal{C}^{1,1}(B'_r)$. \square

Let $\rho : \mathbb{R}^d \rightarrow [0, +\infty)$ be the gauge function of $\mathcal{S}_\delta^\varepsilon$, defined by

$$\rho(x) := \inf\{\lambda > 0 : x \in \lambda \mathcal{S}_\delta^\varepsilon\}, \quad (5.7)$$

where

$$\lambda \mathcal{S}_\delta^\varepsilon := \{\lambda \xi : \xi \in \mathcal{S}_\delta^\varepsilon\}.$$

It is well known that $\rho(\cdot)$ is a positively homogeneous, of degree one, convex function. In particular, we have $\rho(0) = 0$. It is clear to see that $\rho(y) > 0$ at $y \neq 0$.

Lemma 5.9. *The function ρ is locally $\mathcal{C}^{1,1}$ on $\mathbb{R}^d \setminus \{0\}$ and for $y \in \mathbb{R}^d \setminus \{0\}$, $D\rho(y) \in N(y/\rho(y), \mathcal{S}_\delta^\varepsilon)$, where $N(y/\rho(y), \mathcal{S}_\delta^\varepsilon)$ denotes the normal cone at $y/\rho(y) \in \partial \mathcal{S}_\delta^\varepsilon$.*

Proof. It follows from the definition of ρ that for any $y \in \mathbb{R}^d \setminus \{0\}$ and $\lambda > 0$, we have

$$\lambda = \rho(y) \quad \text{if and only if} \quad y/\lambda \in \partial \mathcal{S}_\delta^\varepsilon. \quad (5.8)$$

Fix any $z \in \mathbb{R}^d \setminus \{0\}$ and show that ρ is a $\mathcal{C}^{1,1}$ function near z . Up to an orthogonal change of variables, one can assume that $z = |z|e_d$. It follows from Lemma 5.8 that there exists a neighborhood V of the origin in \mathbb{R}^{d-1} and a function $g \in \mathcal{C}^{1,1}(V)$ such that for any $x' \in V$ and $x_d > 0$,

$$(x', x_d) \in \partial \mathcal{S}_\delta^\varepsilon \quad \text{if and only if} \quad x_d = g(x'). \quad (5.9)$$

Set $\lambda_z = \rho(z)$. By (5.8) and (5.9), we have

$$z_d/\lambda_z = g(z'/\lambda_z) = g(0).$$

Set $I = (\lambda_z/2, 2\lambda_z)$ and $W = (\frac{2}{\lambda_z}V \times (0, \infty))$, and note that $W \times I$ is a neighborhood of $(z, \lambda_z) \in \mathbb{R}^{d+1}$. For any $y = (y', y_d) \in W$, in view of (5.8), we intend to find $\lambda \in I$ such that

$$y/\lambda \in \partial \mathcal{S}_\delta^\varepsilon.$$

Since $y'/\lambda \in V$, the above inclusion reads

$$y_d/\lambda = g(y'/\lambda).$$

We define the function $F : W \times I \rightarrow \mathbb{R}$ by

$$f(y, \lambda) = y_d/\lambda - g(y'/\lambda),$$

note that $f(z, \lambda_z) = 0$, and calculate that

$$\frac{\partial f}{\partial \lambda}(z, \lambda_z) = -\frac{z_d}{\lambda_z^2} < 0.$$

By the implicit function theorem, there are a neighborhood $U \subset W$ of z and a $\mathcal{C}^{1,1}$ function $\Lambda : U \rightarrow I$ such that $f(y, \Lambda(y)) = 0$ for all $y \in U$. In view of (5.8) and (5.9), we see that $\rho(y) = \Lambda(y)$ for all $y \in U$, which completes the proof.

Now, let $y \in \mathbb{R}^d \setminus \{0\}$ and $p = D\rho(y)$. Note that $y/\rho(y) \in \partial\mathcal{S}_\delta^\varepsilon$ and $p = D\rho(y/\rho(y))$ and that for any $x \in \mathcal{S}_\delta^\varepsilon$,

$$1 \geq \rho(x) \geq \rho(y/\rho(y)) + p \cdot (x - y/\rho(y)) = 1 + p \cdot (x - y/\rho(y)),$$

that is, we have

$$p \cdot (x - y/\rho(y)) \leq 0 \quad \text{for all } x \in \mathcal{S}_\delta^\varepsilon.$$

This shows that

$$D\rho(y) = p \in N(y/\rho(y), \mathcal{S}_\delta^\varepsilon). \quad \square$$

Lemma 5.10. *Let $x \in \partial\mathcal{S}_\delta^\varepsilon$ and $i \in \mathcal{I}$. For any $\gamma \in N(x, \mathcal{S}_\delta^\varepsilon)$,*

$$\mathcal{H}_i(\gamma) \begin{cases} \leq 0 & \text{if } x_i < \delta - \varepsilon, \\ \geq 0 & \text{if } x_i > -(\delta - \varepsilon). \end{cases}$$

Proof. Fix any $x \in \partial\mathcal{S}_\delta^\varepsilon$.

Since $\rho(x) = 1$, $\rho(z) \leq 1$ for all $z \in \mathcal{S}_\delta^\varepsilon$, and ρ is convex, we have

$$1 \geq \rho(z) \geq \rho(x) + D\rho(x) \cdot (z - x) \quad \text{for all } z \in \mathcal{S}_\delta^\varepsilon,$$

which implies that $D\rho(x) \in N(x, \mathcal{S}_\delta^\varepsilon)$. Since ρ is differentiable at x and $D\rho(x) \neq 0$, it is easily seen that

$$N(x, \mathcal{S}_\delta^\varepsilon) = \mathbb{R}_+ D\rho(x). \quad (5.10)$$

Set $y := x - \varepsilon D\rho(x)/|D\rho(x)|$ and observe that

$$\bar{B}_\varepsilon(x) \cap \partial\mathcal{S}_\delta = \{y\}, \quad (5.11)$$

which implies that $\bar{B}_\varepsilon(x) \cap \mathcal{S}_\delta = \{y\}$, and therefore

$$D\rho(x) \in N(y, \mathcal{S}_\delta). \quad (5.12)$$

Indeed, it is clear that

$$\bar{B}_\varepsilon(x) \cap \partial\mathcal{S}_\delta \neq \emptyset.$$

Note moreover that if $z \in \bar{B}_\varepsilon(x) \cap \partial\mathcal{S}_\delta$, then $B_\varepsilon(z) \subset \mathcal{S}_\delta^\varepsilon$ and, according to (5.10),

$$D\rho(x) \cdot (w - x) \leq 0 \quad \text{for all } w \in B_\varepsilon(z).$$

This last inequality readily yields

$$z - x = -\varepsilon D\rho(x)/|D\rho(x)|,$$

that is, $z = y$, and we conclude that (5.11) holds.

Fix any $\gamma \in N(x, \mathcal{S}_\delta^\varepsilon)$ and $i \in \mathcal{I}$. By (5.10) and (5.12), we have

$$\gamma = (|\gamma|/|D\rho(x)|)D\rho(x) \in N(y, \mathcal{S}_\delta).$$

Thus, noting that

$$y_i \begin{cases} \leq x_i + |x - y| < \delta & \text{if } x_i < \delta - \varepsilon, \\ \geq x_i - |x - y| > -\delta & \text{if } x_i > -(\delta - \varepsilon), \end{cases}$$

we may apply Lemmas 5.6 and 5.7, to get

$$\mathcal{H}_i(\gamma) \begin{cases} \leq 0 & \text{if } x_i < \delta - \varepsilon, \\ \geq 0 & \text{if } x_i > -(\delta - \varepsilon), \end{cases}$$

which completes the proof. \square

Proof of Theorem 4.2. As noted before, we know that when $d = 1$, the function $\varphi(x) := x^2$ has all the required properties.

In what follows we assume that $d > 1$, and we define

$$\varphi(x) := \rho^2(x) \quad \text{for } x \in \mathbb{R}^d,$$

where ρ is the function given by (5.7). It is clear that $\varphi(x) > 0$ for all $x \in \mathbb{R}^d \setminus \{0\}$ and φ is positively homogeneous of degree two.

Next we show that φ is globally $\mathcal{C}^{1,1}$ on \mathbb{R}^d . Since ρ is locally $\mathcal{C}^{1,1}$ on $\mathbb{R}^d \setminus \{0\}$, the function φ is also locally $\mathcal{C}^{1,1}$ on $\mathbb{R}^d \setminus \{0\}$. Let $x, y \in \mathbb{R}^d \setminus \{0\}$ and in view of the homogeneity property of φ we obtain that

$$\begin{aligned} D\varphi(x) - D\varphi(y) &= |x|D\varphi(x/|x|) - |y|D\varphi(y/|y|) \\ &= |x|(D\varphi(x/|x|) - D\varphi(y/|y|)) + (|x| - |y|)D\varphi(y/|y|). \end{aligned}$$

We choose a constant $C > 0$ so that

$$|D\varphi(p)| \leq C \quad \text{and} \quad |D\varphi(p) - D\varphi(q)| \leq C|p - q| \quad \text{for all } p, q \in \partial B_1.$$

Combining these yields

$$\begin{aligned} |D\varphi(x) - D\varphi(y)| &\leq |x| |D\varphi(x/|x|) - D\varphi(y/|y|)| + (|x| - |y|) |D\varphi(y/|y|)| \\ &\leq C|x| |x/|x| - y/|y|| + C|x - y| \\ &\leq 2C|x - y| + C|x - y| = 3C|x - y|. \end{aligned}$$

This shows that for any $x, y \in \mathbb{R}^d$,

$$|D\varphi(x) - D\varphi(y)| \leq 3C|x - y|.$$

Accordingly, the function φ is globally $\mathcal{C}^{1,1}$ on \mathbb{R}^d .

Let $x \in \mathbb{R}^d \setminus \{0\}$ and $i \in \mathcal{I}$, and assume that $x_i \geq 0$. By Lemma 5.9, we have

$$D\varphi(x) = 2\rho(x)D\rho(x) \in N(x/\rho(x), \mathcal{S}_\delta^\varepsilon).$$

Noting that $x/\rho(x) \in \partial \mathcal{S}_\delta^\varepsilon$, we see from Lemma 5.10 that

$$\mathcal{H}_i(D\varphi(x)) \geq 0.$$

Thus, noting that $D\varphi(0) = 0$ and, hence, $\mathcal{H}_i(D\varphi(0)) = 0$ for all $i \in \mathcal{I}$, we may conclude that for any $x \in \mathbb{R}^d$ and $i \in \mathcal{I}$,

$$\mathcal{H}_i(D\varphi(x)) \geq 0 \quad \text{if } x_i \geq 0.$$

An argument similar to the above shows that for any $x \in \mathbb{R}^d$ and $i \in \mathcal{I}$,

$$\mathcal{H}_i(D\varphi(x)) \leq 0 \quad \text{if } x_i \leq 0.$$

This completes the proof. \square

Remark 5.2. The $\mathcal{C}^{1,1}$ regularity of test functions is sufficient to achieve the comparison theorem as is shown in the proof of Theorem 4.1. Alternatively, one can build smooth test functions by using the mollification techniques.

APPENDIX A.

Let us first recall the Skorokhod problem which will be used in the analysis below.

Definition A.1 (Skorokhod Problem). Given $\psi \in \mathcal{C}([0, \infty) : \mathbb{R})$ with $\psi(0) \geq 0$, a pair $(\phi, \eta) \in (\mathcal{C}([0, \infty) : \mathbb{R}_+))^2$ is said to solve the Skorokhod Problem (SP) for ψ on the domain $[0, \infty)$, if

- (1) $\phi(t) = \psi(t) + \eta(t), \quad t \geq 0,$
- (2) $\phi(0) = \psi(0), \quad t \geq 0,$
- (3) η is non-negative, non-decreasing and $\int_0^t \phi(s) d\eta(s) = 0, \quad \forall t \geq 0.$

In fact, there is always a unique pair (ϕ, η) solving the Skorokhod Problem for a given ψ , given by

$$\eta(t) = \sup_{s \in [0, t]} (-\psi(s))^+, \quad \phi(t) = \psi(t) + \eta(t), \quad t \geq 0.$$

Recall that for $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{R}_+^d$ with $\max_i \alpha_i < 1$, we have

$$\mathbf{M}(\alpha) = \{M = [\mathbb{I} - P] \in \mathbb{R}^d \otimes \mathbb{R}^d : P_{ii} = 0, P_{ij} \geq 0 \text{ and } \sum_{j:j \neq i} P_{ji} \leq \alpha_i\}.$$

In this section we show that for any $v \in \mathfrak{M}$ and given $x \in \mathbb{R}_+^d$, the following reflected diffusion has a unique adapted strong solution $(X, Y) \in \mathcal{C}([0, \infty) : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, \infty) : \mathbb{R}_+^d)$,

$$X(t) = x + \int_0^t b(X(s)) ds + \Sigma W(t) + \int_{[0, t]} v(s) dY(s), \quad t \geq 0, \text{ a.s.}, \quad (\text{A.1})$$

$$\int_{[0, \infty)} X_i(s) dY_i(s) = 0 \quad \text{for all } i \in \mathcal{I},$$

where W is a standard d -dimensional Brownian motion on a given filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, \mathbb{P})$. Here \mathfrak{M} denotes the set of all admissible controls v where v takes values in $\mathbf{M}(\alpha)$. Before we state the result let us define the following quantity,

$$\beta_i(t) = \sup_{0 \leq s \leq t} \max\{0, (-\Sigma W(s))_i\}. \quad (\text{A.2})$$

The following result is a modified version of the results obtained in [14].

Theorem A.1. *For any $v \in \mathfrak{M}$, (A.1) has a unique adapted strong solution (X, Y) taking values in $\mathcal{C}([0, \infty) : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, \infty) : \mathbb{R}_+^d)$ and*

$$\sum_{i \in \mathcal{I}} Y_i(t) \leq \frac{1}{1 - \max_i \alpha_i} \left(|b|_1 t + \sum_{i \in \mathcal{I}} \beta_i(t) \right) \quad \text{for all } t \geq 0, \text{ a.s.}, \quad (\text{A.3})$$

where β is given by (A.2) and $|b|_1 = \sum_{i \in \mathcal{I}} \sup_{x \in \mathbb{R}_+^d} |b_i(x)|$.

We refer, for instance, to [11] for general results for reflected diffusion processes on smooth domains.

Proof. The key ideas of this proof are borrowed from [14]. Fix $T > 0$. To show the existence of a unique adapted strong solution it enough to show the existence of a unique adapted strong solution on the the interval $[0, T]$. Thus we consider the space $\mathcal{C}([0, T] : \mathbb{R}^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$. For $(x, y) \in \mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$, we define

$$\|x\|_T = \sum_{i \in \mathcal{I}} \sup_{s \in [0, T]} |x_i(s)|,$$

$$[y]_T = \sum_{i \in \mathcal{I}} [y_i]_T,$$

where we recall that $[\cdot]_s$ denotes the bounded variation norm on the interval $[0, s]$. For positive γ_1, γ_2 , we endow the space $\mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$ with the metric

$$d((x, y), (\bar{x}, \bar{y})) = \gamma_1 \|x - \bar{x}\|_T + \gamma_2 [y - \bar{y}]_T,$$

where $(x, y), (\bar{x}, \bar{y}) \in \mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$. It is easy to check that for any positive $\gamma_i, i = 1, 2$, $\mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$ forms a complete metric space with the metric d . For $(x, y) \in \mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$ we define two maps $\mathcal{T} : \mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d) \rightarrow \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$ and $\mathcal{S} : \mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d) \rightarrow \mathcal{C}([0, T] : \mathbb{R}_+^d)$ as follows:

$$\mathcal{T}_i(x, y)(t) = \sup_{0 \leq s \leq t} \max\{0, -U_i(x, y)(s)\},$$

$$\mathcal{S}_i(x, y)(t) = U_i(x, y)(t) + \mathcal{T}_i(x, y)(t),$$

$$\text{where } U_i(x, y)(t) = x_i + \int_0^t b_i(x(s))ds + (\Sigma W)_i(t) + \sum_{j \neq i} \int_0^t v_{ij}(s)dy_j(s).$$

From the property of 1-dimensional Skorokhod problem we note that any fixed point of $(\mathcal{S}, \mathcal{T})$ solves (A.1). The idea of the proof is choose suitable γ_i, T_0 so that the map $(\mathcal{S}, \mathcal{T}) : \mathcal{C}([0, T_0] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T_0] : \mathbb{R}_+^d) \rightarrow \mathcal{C}([0, T_0] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T_0] : \mathbb{R}_+^d)$ becomes a contraction almost surely. Once we have uniqueness on the interval $[0, T_0]$ one can extend the result on $[0, T]$ be standard time shifting argument. Now we fix a sample point in $(\Omega, \mathcal{F}, \mathbb{P})$. Then we view X_0, W as deterministic elements. Let K be the Lipschitz constant of b . Let $(x, y), (\bar{x}, \bar{y}) \in \mathcal{C}([0, T] : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, T] : \mathbb{R}_+^d)$. Using the lemma of Shashiasvili [16, pp. 170-175] concerning variational distance between maximal functions, we get

$$\begin{aligned} [\mathcal{T}_i(x, y) - \mathcal{T}_i(\bar{x}, \bar{y})]_T &\leq [U_i(x, y) - U_i(\bar{x}, \bar{y})]_T \\ &\leq \int_0^T |b_i(x(s)) - b_i(\bar{x}(s))|ds + \sum_{j \neq i} \int_0^T (-v_{ij}(s))|dy_j - d\bar{y}_j|(s), \end{aligned}$$

where we have used the property that $v_{ij} \leq 0$. We sum over i to obtain

$$\begin{aligned} [\mathcal{T}(x, y) - \mathcal{T}(\bar{x}, \bar{y})]_T &\leq KdT \|x - \bar{x}\|_T + \sum_{j=1}^d \int_0^T \sum_{i \neq j} (-v_{ij}(s))|dy_j - d\bar{y}_j|(s) \\ &\leq KdT \|x - \bar{x}\|_T + \max_i \alpha_i \sum_{j=1}^d \int_0^T |dy_j - d\bar{y}_j|(s) \end{aligned}$$

$$= KdT \|x - \bar{x}\|_T + \max_i \alpha_i \llbracket y - \bar{y} \rrbracket_T. \quad (\text{A.4})$$

Similarly,

$$\|\mathcal{S}(x, y) - \mathcal{S}(\bar{x}, \bar{y})\|_\infty \leq 2KdT \|x - \bar{x}\|_T + 2 \max_i \alpha_i \llbracket y - \bar{y} \rrbracket_T. \quad (\text{A.5})$$

Therefore combining (A.4) and (A.5) we obtain

$$\begin{aligned} & d((\mathcal{S}(x, y), \mathcal{T}(x, y)), ((\mathcal{S}(\bar{x}, \bar{y}), \mathcal{T}(\bar{x}, \bar{y}))) \\ & \leq KdT \frac{(2\gamma_1 + \gamma_2)}{\gamma_1} \gamma_1 \|x - \bar{x}\|_T + \max_i \alpha_i \frac{(2\gamma_1 + \gamma_2)}{\gamma_2} \gamma_2 \llbracket y - \bar{y} \rrbracket_T. \\ & \leq \max\{KdT \frac{(2\gamma_1 + \gamma_2)}{\gamma_1}, \max_i \alpha_i \frac{(2\gamma_1 + \gamma_2)}{\gamma_2}\} d((x, y), (\bar{x}, \bar{y})). \end{aligned}$$

Now we first choose γ_1, γ_2 to satisfy $\max_i \alpha_i \frac{(2\gamma_1 + \gamma_2)}{\gamma_2} < 1$ (use the fact that $\max_i \alpha_i < 1$) and then choose T small enough to satisfy $KdT \frac{(2\gamma_1 + \gamma_2)}{\gamma_1} < 1$. Hence we have

$$d((\mathcal{S}(x, y), \mathcal{T}(x, y)), ((\mathcal{S}(\bar{x}, \bar{y}), \mathcal{T}(\bar{x}, \bar{y}))) < \varrho d((x, y), (\bar{x}, \bar{y})) \quad \text{for some } \varrho < 1.$$

Therefore one can use contraction mapping theorem to get the unique fixed point in $\mathcal{C}([0, T_0] : \mathbb{R}^d) \times \mathcal{C}_0^\uparrow([0, T_0] : \mathbb{R}^d)$ that solves (A.1).

Now we prove (A.3). Let $(X, Y) \in \mathcal{C}([0, \infty) : \mathbb{R}_+^d) \times \mathcal{C}_0^\uparrow([0, \infty) : \mathbb{R}_+^d)$ be the solution of (A.1). Then

$$Y_i(t) = \sup_{0 \leq s \leq t} \max\{0, -U_i(X, Y)(s)\} \quad \text{for all } i \in \mathcal{I},$$

where

$$U_i(X, Y)(t) = x_i + \int_0^t b_i(X_s) ds + (\Sigma W)_i + \sum_{j \neq i} \int_0^t v_{ij}(s) dY_j(s).$$

Since $x \in \mathbb{R}_+^d$ we have from (A.2)

$$-U_i(X, Y)(t) \leq t \sup_{x \in \mathbb{R}_+^d} |b_i(x)| + \beta_i(t) - \sum_{j \neq i} \int_0^t v_{ij}(s) dY_j(s).$$

Thus using the fact $v_{ij} \leq 0$ for $i \neq j$ we obtain

$$Y_i(t) \leq t \sup_{x \in \mathbb{R}_+^d} |b_i(x)| + \beta_i(t) - \sum_{j \neq i} \int_0^t v_{ij}(s) dY_j(s)$$

Summing over i and using the property that $v(s) \in \mathbf{M}(\alpha)$ we have

$$\sum_{i \in \mathcal{I}} Y_i(t) \leq t |b|_1 + \sum_i \beta_i(t) + \max_i \alpha_i \sum_{i \in \mathcal{I}} Y_i(t),$$

and this proves (A.3). \square

Recall J and V from (2.3) and (2.4). Also recall σ from Proposition 3.1.

Lemma A.1. *Let $\varepsilon > 0$ and $x \in \mathbb{R}^d$. Then for every $v \in \mathfrak{M}$, there exists a control $\tilde{v} \in \mathfrak{M}$ such that*

$$\mathbb{E}_x \left[\int_{\sigma \wedge t}^{\infty} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(\tilde{v}) dY_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \leq e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) + \varepsilon \quad a.s.$$

Proof. Recall the σ_r is the exit time from the open ball $\mathbb{B}_r(x)$. Now using the continuity property of V (Lemma 3.2) and (3.7) we can find $\delta > 0$ so that

$$\sup_{v \in \mathfrak{M}} |J(y, v) - J(\bar{y}, v)| + |V(y) - V(\bar{y})| < \varepsilon \quad \text{for all } y, \bar{y} \in \bar{\mathbb{B}}_r(x), \quad (\text{A.6})$$

and $|y - \bar{y}| \leq \delta$. Now consider a finite Borel partition $\{D_k\}_{k \geq 1}$ of $\bar{\mathbb{B}}_r(x)$ such that $\text{diam}(D_j) < \delta$. Choose $y_k \in D_k$. Then by definition we have $v^k \in \mathfrak{M}$ satisfying

$$J(y_k, v^k) \leq V(y_k) + \varepsilon.$$

Combining with (A.6) we see that for any $y \in D_k$, we have

$$J(y, v^k) \leq J(y_k, v^k) + \varepsilon \leq V(y_k) + 2\varepsilon \leq V(y) + 3\varepsilon. \quad (\text{A.7})$$

Since v^k is adapted to $\{\mathcal{F}_t\}$ we have a progressively measurable map $\Psi_j : [0, \infty) \times \mathcal{C}([0, \infty) : \mathbb{R}^d) \rightarrow \mathbf{M}$ such that

$$v^k(t) = \Psi_k(t, W(\cdot \wedge t)), \quad a.s. \quad \forall t \geq 0.$$

Now for any $v \in \mathfrak{M}$ and the corresponding solution (X, Y) define

$$\tilde{v}(s) = \begin{cases} v(s) & \text{if } s \leq \sigma \wedge t, \\ \Psi_k(s - \sigma \wedge t, W(\cdot \wedge s) - W(\sigma \wedge t)) & \text{if } s > \sigma \wedge t, \text{ and } X(\sigma \wedge t) \in D_k. \end{cases}$$

By strong uniqueness we see that $\sigma \wedge t$ does not change with the new control \tilde{v} . Then denoting $\bar{Y} = Y(\sigma \wedge t + \cdot) - Y(\sigma \wedge t)$, we obtain

$$\begin{aligned} & \mathbb{E}_x \left[\int_{\sigma \wedge t}^{\infty} e^{-\beta s} (\ell(X(s)) ds + \sum_{i \in \mathcal{I}} h_i(\tilde{v}) dY_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \\ &= e^{-\beta \sigma \wedge t} \mathbb{E} \left[\int_0^{\infty} e^{-\beta s} (\ell(X(\sigma \wedge t + s)) ds + \sum_{i \in \mathcal{I}} h_i(\tilde{v}(\sigma \wedge t + s)) d\bar{Y}_i) \middle| \mathcal{F}_{\sigma \wedge t} \right] \\ &= e^{-\beta \sigma \wedge t} J(X(\sigma \wedge t), \tilde{v}(\sigma \wedge t + \cdot)) \\ &\leq e^{-\beta \sigma \wedge t} V(X(\sigma \wedge t)) + 3\varepsilon, \end{aligned}$$

where in the last inequality we use (A.7). \square

ACKNOWLEDGEMENT

Anup Biswas acknowledges the hospitality of the Department of Electrical Engineering (specially, Prof. Rami Atar) in Technion while he was visiting at the early stages of this work. This research of Anup Biswas was supported in part by a INSPIRE faculty fellowship. The research of Hitoshi Ishii was supported in part by the JSPS grants, (KAKENHI #26220702 and #16H03948). The research of Subhamay Saha was supported in part by ISF (grant 1315/12). The research of Lin Wang was supported in part by NSF of China (grant 11401107) and the JSPS grant, (KAKENHI #26220702) and he also would like to thank the second author for the warm and thoughtful invitation to Waseda University.

REFERENCES

- [1] Rami Atar, Amarjit Budhiraja. Singular control with state constraints on unbounded domain. *Ann. Probab.* 34, no. 5, 1864–1909, 2006.
- [2] Rami Atar, Paul Dupuis. A differential game with constrained dynamics and viscosity solutions of a related HJB equation. *Nonlinear Anal.*, no. 7, 1105–1130, 2002.
- [3] Guy Barles. Fully nonlinear Neumann type boundary conditions for second-order elliptic and parabolic equations. *J. Differential Equations* 106, no. 1, 90–106, 1993.
- [4] Anup Biswas. An ergodic control problem for many-server multi-class queueing systems with help. Preprint, 2015.
- [5] Vivek Borkar, Amarjit Budhiraja. Ergodic control for constrained diffusions: characterization using HJB equations. *SIAM J. Control Optim.* 43, no. 4, 1467–1492, (2004/05).
- [6] M. G. Crandall, H. Ishii, P. L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)* 27 1–67, 1992.
- [7] Paul Dupuis, Hitoshi Ishii. On oblique derivative problems for fully nonlinear second order elliptic partial differential equations on nonsmooth domains, *Nonlinear Anal.* 15, No. 12, 1123–1138, 1990.
- [8] Paul Dupuis, Hitoshi Ishii. On oblique derivative problems for fully nonlinear second-order elliptic PDEs on domains with corners. *Hokkaido Math. J.* 20, no. 1, 135–164, 1991.
- [9] Hitoshi Ishii, Moto-Hiko Sato. Nonlinear oblique derivative problems for singular degenerate parabolic equations on a general domain. *Nonlinear Anal.* 57, no. 7-8, 1077–1098, 2004.
- [10] Olav Kallenberg. Foundations of modern probability. Second edition. Probability and its Applications (New York). Springer-Verlag, New York, 2002.
- [11] P.-L. Lions and A.-S. Sznitman, Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.* 37, no. 45, 11–537, 1984.
- [12] Kavita Ramanan and Martin I. Reiman. Fluid and heavy traffic diffusion limits for a generalized processor sharing model. *Ann. Appl. Probab.* 13, no. 1, 100–139, 2003.
- [13] Kavita Ramanan and Martin I. Reiman. The heavy traffic limit of an unbalanced generalized processor sharing model. *Ann. Appl. Probab.* 18, no. 1, 22–58, 2008.
- [14] S. Ramasubramanian. A subsidy-surplus model and the Skorokhod problem in an orthant. *Math. Oper. Res.* 25, no. 3, 509–538, 2000.
- [15] R. Tyrrell Rockafellar. Convex analysis, Princeton Mathematical Series, No. 28 Princeton University Press, Princeton, N.J. 1970.
- [16] M. Shashivili. A lemma of variational distance between maximal functions with application to the Skorokhod problem in a nonnegative orthant with state dependent reflection directions. *Stochast. Rep.* 48, no. 3-4, 161–194, 1994.
- [17] Jiongmin Yong, Xun Yu Zhou. Stochastic controls. Hamiltonian systems and HJB equations, Applications of Mathematics (New York), 43. Springer-Verlag, New York, 1999.

DEPARTMENT OF MATHEMATICS, INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH, DR. HOMI BHABHA ROAD, PUNE 411008, INDIA.

E-mail address: `anup@iiserpune.ac.in`

FACULTY OF EDUCATION AND INTEGRATED ARTS AND SCIENCES, WASEDA UNIVERSITY, SHINJUKU, TOKYO, 169-8050 JAPAN.

E-mail address: `hitoshi.ishii@waseda.jp`

DEPARTMENT OF ELECTRICAL ENGINEERING, TECHNION, HAIFA 32000, ISRAEL.

E-mail address: `subhamay585@gmail.com`

YAU MATHEMATICAL SCIENCES CENTER, TSINGHUA UNIVERSITY, BEIJING, 100084, CHINA.

E-mail address: `lwang@math.tsinghua.edu.cn`